

**SILVIA FIGUEIREDO BRANDÃO**

Organizadora

**DUAS VARIEDADES AFRICANAS DO PORTUGUÊS:  
VARIÁVEIS FONÉTICO-FONOLÓGICAS  
E MORFOSSINTÁTICAS**



**Blucher** Open Access



DUAS VARIEDADES AFRICANAS  
DO PORTUGUÊS:  
VARIÁVEIS FONÉTICO-FONOLÓGICAS  
E MORFOSSINTÁTICAS

2018

**Blucher**

Duas variedades africanas do Português: variáveis fonético-fonológicas e morfossintáticas

© 2018 Silvia Figueiredo Brandão

Editora Edgard Blücher Ltda.

*As imagens utilizadas nesta obra são de responsabilidade dos autores e todas contam com as devidas referências de onde foram extraídas.*

---

# Blucher

---

Rua Pedroso Alvarenga, 1245, 4º andar

04531-934 – São Paulo – SP – Brasil

Tel 55 11 3078-5366

**contato@blucher.com.br**

**www.blucher.com.br**

Segundo Novo Acordo Ortográfico, conforme 5. ed.  
do *Vocabulário Ortográfico da Língua Portuguesa*,  
Academia Brasileira de Letras, março de 2009.

É proibida a reprodução total ou parcial por  
quaisquer meios, sem autorização escrita da Editora.

---

Todos os direitos reservados pela Editora  
Edgard Blücher Ltda.

Dados Internacionais de Catalogação na Publicação (CIP)  
Angélica Ilacqua CRB-8/7057

---

Duas variedades africanas do português : variáveis  
fonético – fonológicas e morfossintáticas / Silvia Figueiredo  
Brandão (org.). -- São Paulo : Blucher, 2018.

326 p.

Bibliografia

ISBN 978-85-8039-325-5 (impresso)

ISBN 978-85-8039-324-8 (e-book)

1. Língua portuguesa – Brasil 2. Língua portuguesa –  
São Tomé e Príncipe 3. Língua portuguesa – Moçambique  
4. Língua Portuguesa – Fonética - Variação 5. Linguística I.  
Brandão, Sílvia Figueiredo

18-0806

CDD 469.7

---

Índices para catálogo sistemático:

1. Língua portuguesa – Variações

DUAS VARIEDADES AFRICANAS  
DO PORTUGUÊS:  
VARIÁVEIS FONÉTICO-FONOLÓGICAS  
E MORFOSSINTÁTICAS



# SUMÁRIO

|                   |    |
|-------------------|----|
| Prefácio .....    | 9  |
| Apresentação..... | 11 |

## I

### PERFIL DAS COMUNIDADES

|  |  |           |
|--|--|-----------|
| CAPÍTULO 1   |  |           |
| <b>A questão do contato linguístico .....</b>  |  | <b>21</b> |
| <i>Fabiane de Mello Vianna da Rocha Teixeira Rodrigues do Nascimento</i>               |  |           |
| CAPÍTULO 2   |  |           |
| <b>São Tomé e Príncipe: aspectos históricos, econômico-sociais e linguísticos.....</b> |  | <b>43</b> |
| <i>Fabiane de Mello Vianna da Rocha Teixeira Rodrigues do Nascimento</i>               |  |           |
| CAPÍTULO 3   |  |           |
| <b>O perfil multilíngue de Moçambique.....</b>   |  | <b>75</b> |
| <i>Karen Cristina da Silva Pissurno</i>  |  |           |

## II

### VARIÁVEIS FONÉTICO-FONOLÓGICAS

|  |  |            |
|--|--|------------|
| CAPÍTULO 4   |  |            |
| <b>Róticos nas variedades urbanas santomense e moçambicana do Português .....</b>    |  | <b>95</b>  |
| <i>Silvia Figueiredo Brandão e Alessandra de Paula</i>                               |  |            |
| CAPÍTULO 5   |  |            |
| <b>As vogais médias pretônicas na variedade urbana do Português de São Tomé.....</b> |  | <b>119</b> |
| <i>Fabiane de Mello Vianna da Rocha Teixeira Rodrigues do Nascimento</i>             |  |            |
| CAPÍTULO 6   |  |            |
| <b>Vogais em contexto postônico medial no Português de São Tomé.....</b>             |  | <b>159</b> |
| <i>Danielle Kely Gomes</i>   |  |            |
| CAPÍTULO 7   |  |            |
| <b>O ditongo /ei/ na fala de São Tomé .....</b>                                      |  | <b>177</b> |
| <i>Raphaella Ribeiro Passos</i>  |  |            |

III  
VARIÁVEIS MORFOSSINTÁTICAS

CAPÍTULO 8

**Concordância nominal no Português de São Tomé e no Português de Moçambique** ..... 203

*Silvia Figueiredo Brandão*

CAPÍTULO 9

**Padrões de concordância verbal de terceira pessoa plural no Português de São Tomé e no Português de Moçambique** ..... 245

*Karen Cristina da Silva Pissurno*

CAPÍTULO 10

**Ordem dos clíticos pronominais no Português de São Tomé e no Português de Moçambique** ..... 277

*Silvia Rodrigues Vieira e Maria de Fátima Vieira*

# PREFÁCIO

Foi com grande satisfação que, em janeiro de 2018, nas instalações do Centro de Linguística da Universidade de Lisboa (CLUL), recebi, de viva voz, o convite que a minha estimada colega, a Professora Sílvia Brandão, me endereçou para prefaciara presente obra, que irá, sem dúvida, contribuir para enriquecer a produção e o conhecimento sobre o português falado em África, mais precisamente em Moçambique e São Tomé e Príncipe, na sua relação com a variedade brasileira. O conjunto de capítulos que compõe este trabalho tem como propósito investigar a que se devem determinadas tendências linguísticas de convergência e divergência no espaço da língua portuguesa onde o português não é ou não foi, historicamente, a língua dominante. Trata-se de uma obra com assinatura exclusiva da Universidade Federal do Rio de Janeiro e integralmente no feminino (o prefaciador será a exceção que confirma a regra!), para a qual também contribuiu, modestamente, o Centro de Linguística da Universidade de Lisboa, através de diversos tipos de colaboração e partilha de recursos com a URFJ, o que mostra a unidade na diversidade em matéria de assuntos da língua portuguesa.

A investigação científica sobre o português em África é um fenómeno essencialmente pós-colonial que se encontra em franca expansão devido a uma conjuntura favorável que se foi desenvolvendo ao longo das últimas décadas. Neste percurso, devemos destacar, acima de tudo, a conquista da independência por parte das antigas colónias portuguesas em África, a qual funcionou como um primeiro catalisador de descentralização linguística, ainda que o português europeu tenha sido elevado a língua oficial exclusiva nesses novos Estados. O fim dos longos conflitos armados que se seguiram às independências em Angola e Moçambique foi um outro passo indispensável para que se criassem condições propícias à investigação. Paradoxalmente, estes conflitos tiveram como “efeito secundário” uma maior propagação do português, visto que muitos falantes de outras línguas, oriundos de zonas interiores, foram forçados a abrigar-se nos centros urbanos, onde o português funciona(va) como língua franca. A estabilidade alcançada depois deste período conturbado continua a favorecer o português, a principal língua da educação, da administração, dos meios de comunicação e da ascensão socioeconómica, isto é, a língua alta.

Nos últimos anos, graças a uma maior estabilidade, tem-se vindo a realçar a crescente importância económica e geolinguística dos países de língua portuguesa em África, especialmente Angola e Moçambique. Nestes países, o elevado cres-

cimento demográfico, que se faz acompanhar de um aumento muito significativo do número de falantes L1 e L2 do português, constitui um pilar fundamental na manutenção do português como uma das línguas com relevância global e a mais falada no hemisfério sul. O processo de crescente nativização do português em África é, de resto, um fenómeno singular que não se observa em relação a nenhuma outra antiga língua colonial nesse continente, sendo São Tomé e Príncipe o caso mais emblemático, onde, salvo raríssimas exceções, todos falam português e predominantemente como L1. Este percurso independente, mas sempre na companhia do português, por parte dos países africanos de língua oficial portuguesa traduz-se, inevitavelmente, num crescendo de autonomia linguística, face às normas do PE e do PB, que se foram construindo há mais tempo. A afirmação destas novas variedades significa, pois, um gradual abandono do modelo bicêntrico e bicontinental em que se apoiava a língua portuguesa em direção a um modelo pluricêntrico e tricontinental. São ecos do grito de Ipiranga que continuam a ressoar. A bem da língua portuguesa, como não podia deixar de ser.

A delimitação do objeto de estudo da presente obra, estudos variacionistas sobre aspetos fonético-fonológicos e morfossintáticos das variedades de português de Moçambique e S. Tomé e Príncipe, configura uma escolha feliz de vários pontos de vista. Em primeiro lugar, trata-se de duas variedades com características sociolinguísticas distintas: a primeira está em contacto com línguas aglutinantes do grupo banto e apresenta uma taxa de nativização crescente, embora ainda relativamente baixa; a segunda constitui a L1 da maioria da população, mesmo que os censos nacionais não explicitem esta informação. Neste caso, o contacto com o crioulo de São Tomé, uma língua de tipologia isolante, condicionou o português sobretudo historicamente. À comparação entre estas duas variedades junta-se uma terceira, o português de Brasil, uma L1 consolidada com forte historial de L2, com um papel de relevo para o continente africano que se traduz na hipótese de um *continuum* afro-brasileiro. O contraste entre estas três variedades promete dar pistas mais robustas na identificação dos fatores que determinam a variação e mudança linguísticas. Segundo, é uma grande mais-valia a inclusão de estudos detalhados de natureza fonético-fonológica, visto tratar-se de uma área que tem sido largamente descurada na investigação sobre o português em África. Em terceiro lugar, a opção metodológica pela abordagem variacionista não só reforça ainda mais a consistência desta obra como também se afasta de uma certa tradição que apenas privilegia a descrição dos fenómenos divergentes, ignorando, em larga medida, a variação, inclusive a convergência com a norma, e a respetiva interpretação. De facto, a quantificação dos fenómenos relativamente a um *corpus* linguístico pré-definido constitui uma abordagem particularmente valiosa para a descrição e análise de contextos linguísticos mais instáveis e permite formular novas questões sobre gramáticas em competição e processos de cristalização das diferentes áreas da gramática em variedades jovens.

Termino este prefácio felicitando a equipa de editoras e autoras da UFRJ e na expectativa de que esta obra sirva de inspiração a novos trabalhos que possam continuar a documentar e investigar a evolução das novas variedades, tanto na sua relação com o PE e o PB e com as outras línguas faladas nesses espaços como na identificação dos mecanismos que induzem a mudança linguística. Espera-se também que a investigação sobre o contexto linguístico africano possa, em última instância, culminar em políticas linguísticas mais inclusivas que visem, por um lado, adequar o processo de ensino e aprendizagem do português à complexa realidade (socio)linguística destes espaços e, por outro lado, valorizar as outras línguas, muitas das quais fragilizadas pelas dinâmicas linguísticas em curso, em parceria com investigadores e agentes educativos locais. Só assim se completa o ciclo que começa e termina naqueles falantes anónimos que são a fonte do conhecimento produzido.

Lisboa, fevereiro de 2018.

**TJERK HAGEMEIJER**

Faculdade de Letras da Universidade de Lisboa/  
Centro de Linguística da Universidade de Lisboa



# APRESENTAÇÃO

Nesta obra, reúnem-se estudos desenvolvidos por professoras pesquisadoras vinculadas à Universidade Federal do Rio de Janeiro-UFRJ e que têm por objetivo observar variáveis fonético-fonológicas e morfossintáticas no âmbito do Português de São Tomé (PST) e do Português de Moçambique (PM), de modo a verificar possíveis convergências e divergências entre essas variedades e o Português do Brasil (PB).

São quatro os principais projetos que os originaram: *O vocalismo átono no Português do Brasil e no Português de São Tomé*, apoiado pela FAPERJ<sup>1</sup>, *Três variedades do Português em contraste*, apoiado pelo CNPq<sup>2</sup>, ambos coordenados por Silvia Figueiredo Brandão, *Estudo comparado dos padrões de concordância em variedades africanas, brasileiras e europeias do Português*, o Projeto 21 da ALFAL<sup>3</sup>, coordenado pela referida pesquisadora e por Silvia Rodrigues Vieira, que também coordena o Projeto *Padrões de concordância em variedades africanas, brasileiras e europeias do português: a natureza das restrições e o contato linguístico*<sup>4</sup>.

Nortearam a consecução de todos eles, fundamentalmente, os pressupostos teórico-metodológicos da Teoria da Variação e Mudança (WEINREICH; LABOV; HERZOG, 1968), bem como os desdobramentos da Sociolinguística Variacionista representados nas obras de Labov (1972, 1994, 2001, 2003, entre outras). Os estudos aqui presentes também baseiam-se em *corpora* comuns. As amostras do Português de São Tomé foram selecionadas do *Corpus* VAPOR, do Centro de Linguística da Universidade de Lisboa (CLUL-PT), e as do Português de Moçambique, do *Corpus* Moçambique do Projeto *Estudo comparado dos padrões de concordância em variedades africanas, brasileiras e europeias do Português* (UFRJ), acima referido.

De modo a contribuir para os debates sobre as origens do PB, partiu-se da hipótese de que, dado o caráter multilíngue das comunidades santomense e mo-

---

<sup>1</sup> Bolsa Cientista do nosso Estado – Processo E-26/201.435/2014.

<sup>2</sup> Bolsa de Produtividade em Pesquisa – Processo 304038/2016-4.

<sup>3</sup> Cf. <http://www.mundoalfal.org/sites/default/files/proyectos/EstudoSS.htm>. Cf. também o site [www.concordancia.lettras.ufrj.br](http://www.concordancia.lettras.ufrj.br).

<sup>4</sup> Bolsa Cientista do nosso Estado – Processo E-26/201.436/2014.

çambicana, seria possível nelas encontrar pistas sobre o que teria ocorrido com o Português nos períodos em que, no Brasil, ele coexistia, de forma mais efetiva, com línguas africanas e indígenas. Tal hipótese ancora-se no princípio do Uniformitarismo, proposto por Labov (1972) e segundo o qual forças que concorrem, no presente, para a variação e a mudança linguísticas são do mesmo tipo das que operaram no passado.

Em última instância, o que se pretende, no futuro, é traçar, com base em diferentes variáveis, um *continuum* em que se possa representar o maior ou menor distanciamento das variedades não europeias em relação ao Português Europeu (PE), aquilatar os efeitos do contato multilinguístico que parece determinar as especificidades das variedades africanas e brasileira e, conseqüentemente, discutir questões referentes à formação do PB.

Apesar de a questão da influência no PB das línguas indígenas autóctones e das línguas africanas transplantadas para o Brasil em decorrência da política escravocrata que dominou os primeiros anos da colonização já estar presente, entre outras, nas obras de filólogos brasileiros, como Serafim da Silva Neto (1950), a ideia da existência de um *continuum* afro-brasileiro vem sendo modernamente desenvolvida sobretudo nos trabalhos de Petter (2007, 2015) e de Avelar; Galves (2014, entre outros) sob diferentes prismas teóricos.

Petter (2015: 315), cujos trabalhos se concentram sobretudo no Português de Angola, afirma que “o estudo das variedades faladas na África desfruta de uma situação privilegiada, pois as línguas africanas ainda são faladas e interagem com o português. É interessante observar essa situação, pois ela pode fornecer dados cruciais para o entendimento dos contatos linguísticos e dos processos de mudança em toda a área de expansão da língua portuguesa”.

Na mesma linha, Avelar; Galves (2014: 243) sublinham que seus estudos “advogam em favor da hipótese do contato, explorando a ideia de que certas marcas gramaticais singularizadoras do português brasileiro no conjunto das línguas românicas se devem à ação dos contatos interlinguísticos estabelecidos entre falantes de português e de línguas africanas (em particular, línguas bantas)”.

Em virtude, justamente, de se acreditar ser possível contribuir para um melhor conhecimento do PB por meio do que se verifica hoje nos países africanos de Língua Portuguesa, decidiu-se iniciar as pesquisas observando duas variedades com características sociais e linguísticas específicas.

Brandão (2016: 69-70) indica as motivações que levaram à escolha do Português de São Tomé e Príncipe:

- (a) dentre os países africanos que têm o Português como língua oficial, é a única variedade falada como L1 ou L2 pela maioria da população – 98,4%, em 2012<sup>5</sup> – sendo, hoje, a L1 da maior parte dela;
- (b) em São Tomé e Príncipe, coexistem, além de quatro crioulos de base portuguesa – o Forro (ou Santomé) e o Angolar, 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).
- (c) os quatro primeiros crioulos citados, que, segundo Hagemeyer (2009: 4), hoje “línguas distintas, tiveram uma história comum”, partilham, como informa Ferraz (1976: 9), uma série de características e têm como substrato línguas maternas também comuns às dos africanos trazidos para o Brasil para servirem de escravos, o que amplia o interesse em observá-las em contraste com o PB;
- (d) o Forro, em especial – usado, segundo o censo de 2001, por 72,4% da população mas reduzido, hoje, segundo o Censo de 2012, a apenas 36,3% de falantes – é considerado pelos habitantes da ilha de São Tomé como a “língua nacional”, embora, por ser estigmatizado, nem todos o dominem ou o utilizem nas situações intercomunicativas cotidianas;
- (e) nas escolas, a norma de referência é o Português Europeu, conquanto, pelo que se observa, inclusive na fala de professores nativos, de maior nível de escolaridade, ela nem sempre seja observada no que respeita aos planos sintático, morfossintático e fonológico, conforme consta da observação de Hagemeyer (2009: 19) acima citada.

Já a escolha da variedade urbana do PM, representada por Maputo, leva em conta (GONÇALVES, 2010: 26-35) que,

- (a) contrariamente ao que ocorre em São Tomé, o quadro de falantes do Português é mais reduzido: apenas 10,5% da população o têm como L1, embora seja ela a língua oficial;
- (b) “a comunidade moçambicana de falantes de Português constituiu-se muito recentemente, tendo sido praticamente nula a difusão desta língua durante os primeiros quatro séculos de colonização”: só em meados do século XX, com a chegada de 140 000 colonos deu-se a “colonização maciça” desse território;
- (c) coexistem com o Português, de estatuto minoritário, diversas línguas Banto, também minoritárias, entre as quais o Macua, o Changana, o Lomwe, o Sena e o Chuabo;

---

<sup>5</sup> Instituto Nacional de Estatística (<http://www.ine.st/demografia.html>).

(d) a grande maioria dos falantes de Português concentra-se nas áreas urbanas (72,4%) pelo fato de nelas existir a maior proporção de indivíduos com educação formal;

(e) os falantes de Português L1 predominam entre os mais jovens, decrescendo seu número entre os de mais de 50 anos.

Segundo, ainda, Gonçalves (2010), “uma razão importante para a fraca difusão do português em Moçambique reside na definição de uma política educacional para as colônias portuguesas” (p. 29), que só teria ocorrido a partir de 1930. Efetivamente, “só na década de 60-70 se registou um crescimento considerável de escolas dos vários graus de ensino (primário, secundário e técnico-profissional)” (p. 30). A partir de então, “o crescimento espetacular da população escolar e o alargamento dos contextos de utilização do português traduziram-se naturalmente por um aumento significativo do número de falantes dessa língua” (p. 32).

Uma visão mais abrangente sobre as comunidades alvo das pesquisas que ensejaram o estudo das variáveis aqui focalizadas encontra-se na Seção I do livro, composta por três capítulos.

No capítulo 1, Fabiane Nascimento focaliza questões relacionadas ao contato linguístico, que servirão de base para a melhor compreensão do que ocorre nas duas áreas e, ainda, para o debate sobre a gênese do PB.

O capítulo 2, da mesma autora, constitui uma síntese não só do processo de colonização de São Tomé e Príncipe, com ênfase nos diferentes ciclos histórico-econômicos do país, mas também uma apreciação sobre as diferentes línguas que nele coexistem.

No capítulo 3, Karen Cristina da Silva Pissurno, com base em literatura especializada e em sua experiência na pesquisa de campo em Maputo que redundou na formação do já citado *Corpus* Moçambique, traça um panorama da complexa realidade linguística e social desse país. Os três textos, apesar de seu caráter introdutório, são de fundamental importância para que se possam aquilatar os resultados expostos nos demais estudos.

A Seção II, constituída por quatro capítulos, está dedicada a variáveis fonético-fonológicas.

No capítulo 4, Silvia Figueiredo Brandão e Alessandra de Paula apresentam um estudo sobre os róticos, observando-os em contextos tanto pré-vocálicos quanto pós-vocálicos. As autoras concluem que, no PST e no PM, predominam as variantes mais conservadoras dos róticos (o tepe e a vibrante alveolar) e que, ao que tudo indica, neles não há a consciência fonológica da oposição entre um R [+ant] e um R [-ant] presente no PB e no PE.

No capítulo 5, Fabiane Nascimento estuda as vogais médias pretônicas no PST, buscando, em particular, determinar os fatores que concorrem para o alteamento. Seus resultados sugerem que, embora a norma de referência seja a europeia, os baixos *inputs* da regra de elevação (/e/: .40; /o/: .34) e a manifestação do processo de harmonização vocálica, na série tanto anterior quanto posterior, parecem confirmar a hipótese de proximidade entre crioulos de base portuguesa, PST e PB, no que toca à insubmissão à regra geral de redução, conforme Marquilhas (2003).

No capítulo 6, Danielle Kely Gomes focaliza as vogais postônicas mediais no PST e no PM, observando as restrições que determinam não só o seu apagamento, mas também o alteamento das vogais médias.

No capítulo 7, Raphaela Ribeiro Passos ocupa-se do ditongo /ei/ em contexto medial de vocábulo, como em *beijo*, e em contexto final em formas verbais, como em *cheguei*. Suas análises demonstram não haver qualquer registro da concretização de /ei/ como [ɛj], a exemplo do que ocorre na norma lisboeta do PE, e estar a monotongação em contexto medial sujeita aos mesmos condicionamentos que se observam no PB. Já em contexto final, esse processo está presente sobretudo na fala de indivíduos que utilizam o Forro com maior frequência.

A seção III, formada por três capítulos, diz respeito a variáveis morfossintáticas.

No capítulo 8, Silvia Figueiredo Brandão examina a concordância nominal, fazendo uma síntese dos estudos por ela realizados no âmbito do PST e apresentando a questão também no PM, concluindo que, nas duas variedades, a regra é variável e está sujeita a condicionamentos sociais e estruturais como se verifica no PB.

No capítulo 9, Karen Cristina da Silva Pissurno revisita trabalhos sobre a concordância verbal de terceira pessoa do plural no PST realizados no âmbito do mencionado Projeto da ALFAL e se ocupa, em especial, de pesquisa sobre o tema no PM que redundou em sua dissertação de mestrado.

No capítulo 10, Silvia Rodrigues Vieira e Maria de Fátima Vieira tratam da ordem dos clíticos no PST e no PM, buscando estabelecer padrões convergentes e/ou divergentes em relação aos que se observam no PB e no PE.

Com este conjunto de estudos, espera-se não só contribuir para o maior conhecimento do Português, mas também incentivar outros pesquisadores a empreenderem novas pesquisas sobre as variedades africanas em diferentes enfoques teórico-metodológicos, de modo que se possa, daqui a algum tempo, traçar o tão almejado *continuum* das variedades do Português.

Por fim, cabe passar aos agradecimentos:

- ao Prof. Dr. Tjerk Hagemeyer, renomado crioulista, que não só disponibilizou à equipe os *corpora* representativos do Português de São Tomé, mas também concedeu a honra de prefaciar este livro;
- às colegas autoras, que compartilham o interesse pelo estudo das variedades do Português, por, prontamente, aceitarem o convite para participar desta publicação;
- aos alunos de graduação engajados, com ou sem bolsa, em atividades de Iniciação Científica e/ou em trabalhos de conclusão de curso nos diferentes projetos que deram origem a este livro: Davi Bretas dos Santos Pessanha (Bolsa PIBIC/CNPq), Stefany de Paula Santos (Bolsa de IC/FAPERJ), Helen Lorena Rodrigues Elias Cordeiro (Bolsa PIBIC/UFRJ), Sofia dos Santos Alves (PIBIC/UFRJ), Thalles Candal Reis Fernandes (IC/FAPERJ), Monique Oliveira Corrêa, Nayse Hevellyn Magalhães Barcelos, Mateus Almeida do Pranto, Laura Cunha Calzolari, Amanda Carolina de Oliveira Santos, Maria de Fátima Vieira; e
- à FAPERJ, que, por meio da Bolsa Cientista do nosso Estado (2015-2017), concedida à organizadora da obra, tornou possível esta publicação.

Rio de Janeiro, abril de 2018.

**SILVIA FIGUEIREDO BRANDÃO**  
(Organizadora)

## REFERÊNCIAS

- AVELAR, Juanito; GALVES, Charlotte. O papel das línguas africanas na emergência da gramática do português brasileiro. *Linguística*, 30 (2): 241-288, 2014.
- BRANDÃO, Silvia Figueiredo. Aspectos da variedade urbana do Português de São Tomé: resultados e metas de pesquisa. In: AGUILERA, V. A.; DOIRON, M. P. B. (Org.). *Estudos geossociolinguísticos brasileiros e europeus: uma homenagem a Michel Contini*. Cascavel-PR: EDUNIOESTE; Londrina-PR: EDUEL, 2016. p. 67-87.
- FERRAZ, Luiz Ivens. *The creole of São Tomé*. Johannesburg: Witwatersrand University Press, 1979.
- GONÇALVES, Perpétua. *A gênese do Português de Moçambique*. Lisboa: Imprensa Nacional/Casa da Moeda, 2010.
- HAGEMEIJER, Tjerk. As línguas de São Tomé e Príncipe. *Revista de Crioulos de Base Lexical Portuguesa e Espanhola*, Macau, 1(1): 1-27, 2009.

- LABOV, William. *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press, 1972.
- LABOV, William. *Principles of linguistic change: internal factors*. Oxford: Blackwell Publishers, 1994. v. 1.
- LABOV, William. *Principles of linguistic change: social factors*. Oxford: Blackwell Publishers, 2001. v. 2.
- LABOV, William. Some sociolinguistic principles. In: PAULSTON, C. B.; TUCKER, G. R. (Ed.). *Sociolinguistics: the essential readings*. Blackwell Publishing, 2003. p. 234-250.
- MARQUILHAS, Rita. Mudança analógica e elevação das vogais pretónicas. In: CASTRO, I.; DUARTE, I. (Eds.). *Razões e emoção*. Miscelânea de estudos em homenagem a Maria Helena Mira Mateus. Lisboa: Imprensa Nacional-Casa da Moeda, 2003. p. 7-18.
- PETTER, Margarida Maria Tadonni. Ampliando a investigação do *continuum* afro-brasileiro de português. *Papia*, São Paulo, 25(2): 305-317, 2015.
- WEINREICH, Uriel; LABOV, William; HERZOG, Marvin. Empirical foundations for a theory of linguistic change. In: LEHMANN, W.; MALKIEL, Y. (Org.). *Directions for historical linguistics*. Austin: University of Texas Press, 1968, p. 97-195.



I

# PERFIL DAS COMUNIDADES



# A QUESTÃO DO CONTATO LINGUÍSTICO

**FABIANE DE MELLO VIANNA DA ROCHA  
TEIXEIRA RODRIGUES DO NASCIMENTO**

## **1. INTRODUÇÃO**

A linguagem é uma espécie de comportamento social, uma das maneiras empregadas pelos indivíduos para interagir (cf. HOLM, 2004: 1). Como fato histórico, as línguas surgem, convivem, desaparecem, reinventam-se e funcionam em condições específicas (cf. ORLANDI, 2014: 211). Elas se submetem a processos naturais de mudança, mas os fatores relativos à sua atuação lhes conferem uma complexidade, acentuada na investigação de regiões multilíngues.

Reconhecendo a relevância do contato interlinguístico para o tratamento das variedades do Português faladas em África, este capítulo discute a questão, concentrando-se, sobretudo, no processo de criouliização. Mais especificamente, em um primeiro momento, expõem-se algumas informações sobre a origem do campo científico em destaque (seção 2). A seção 3 concentra-se em aspectos sociolinguísticos comuns aos contextos em que línguas crioulas emergem, ressaltando os pontos linguísticos (3.1) e sócio-históricos (cf. seção 3. 2) convergentes. Encerram o capítulo as principais hipóteses que norteiam a crioulogênese (seção 4) e algumas considerações finais (seção 5).

## 2. A EMERGÊNCIA CIENTÍFICA DO CONTATO LINGUÍSTICO

O contato linguístico<sup>1</sup> suscitou interesse muito antes do século XIX, quando emergiram os primeiros tratados científicos sobre linguagem<sup>2</sup>, mas destacou-se, naquele momento, no âmbito da linguística histórica, pois diversos trabalhos abordaram a questão ao debaterem a natureza das mudanças (cf. SCHMIDT, 1872; MULLER, 1861; SCHUCHARDT, 1882, entre outros). A maneira como o convívio entre falares afeta a gramática das línguas era obscura e o modelo tradicional histórico-comparativo a explicava, distinguindo situações de transmissão regular e irregular (cf. THOMASON; KAUFMAN, 1988: 11; LUCCHESI; BAXTER; RIBEIRO, 2009). Segundo Lucchesi; Baxter; Ribeiro (2009: 101), esta última designa processos históricos de intenso contato interétnico, ocorridos entre os séculos XVI e XIX, quando, em virtude do colonialismo europeu em diferentes partes do mundo (África, Ásia e América), falantes de línguas distintas dividiram o mesmo território. Era comum a língua de superstrato, falada pelo grupo dominante, se tornar alvo da população adulta subjugada e marginalizada. O segundo idioma, aprendido de modo precário e defectivo, servia de modelo para a construção da língua materna das futuras gerações, à proporção que as camadas dominadas se distanciavam de suas origens. Trata-se, pois, de um processo irregular de nativização da língua-alvo, pois o *input* disponível aos falantes nativos da nova sociedade é repleto de lacunas e de reanálises gramaticais, correspondendo a “versões de segunda língua desenvolvidas entre os falantes adultos”. Em contrapartida, na chamada transmissão geracional regular, os dados linguísticos a que a criança se expõe são extraídos da língua materna de seus pais, sendo, portanto, mais completos. Assim, pode-se dizer que, sem descartar a interferência de fatores externos no passado, a difusão regular (cf. MULLER, 1861; MEILLET, 1921; OKSAAR, 1972a, entre outros) se pauta na reconstrução de uma só afiliação genética; envolve, majoritariamente, uma única fonte; e altera a gramática de modo paulatino. Ressalta a impenetrabilidade estrutural, ou seja, a rara ou nula combinação de vernáculos e a evolução dos sistemas, desde uma origem singular e o desenvolvimento de fatores puramente internos ao longo do tempo (cf. WINFORD, 2003: 7). Em contrapartida, na transmissão irregular (cf. WHITNEY, 1881; SCHUCHARDT, 1882, entre outros) salienta-se a legitimidade da mistura (cf. WINFORD, 2003: 6-7), viabilizando a “formação de uma língua historicamente nova, denominada

---

<sup>1</sup> Este capítulo foi integralmente reproduzido da tese de Nascimento (2018).

<sup>2</sup> De acordo com Winford (2003: 6), a temática era frequente entre estudantes de línguas muito antes do século XIX. Por exemplo, em 1666, G. Lucio discute a mistura de dialetos croatas e românicos na Dalmácia, com base em registros do século XIV (cf. SCHUCHARDT, 1884: 30 *apud* WINFORD, 2003: 6).

“língua crioula”, ou a simples formação de uma nova variedade histórica da língua de superstrato, que não deixa de apresentar processos de variação e mudança induzidos pelo contato”<sup>3</sup> (LUCCHESI; BAXTER; RIBEIRO, 2009: 101).

Paralelamente aos debates apontados, desenvolveu-se uma consistente tradição de pesquisa sobre o assunto. Em 1928, Trubetzkoy definiu o conceito de “união de línguas” ou “área linguística” (“Sprachbund”) e iniciou os estudos sobre o tema ao redor do mundo. A maioria dos tratados considerava i) áreas geográficas específicas; ii) processos linguísticos, tipos de alteração e de restrição motivados pelo contato entre falares diversos; iii) empréstimos lexicais; iv) a influência das variedades de substrato na mudança; além de v) casos particulares de línguas mistas, relativos à comutação de códigos bilíngues ou à formação de *pidgins* e de crioulos<sup>4</sup> (cf. WINFORD, 2003: 7-8). Comprovando a importância de aliar componentes estruturais e socioculturais, às postulações assinaladas somaram-se informações a respeito do contexto social em que o contato se efetiva. As contribuições de áreas como a antropologia, a sociologia da linguagem, a sociolinguística e a psicologia conferiram-lhe um *status* interdisciplinar decisivo na sua promoção a campo teórico. Finalmente, em junho de 1979, o Primeiro Congresso Mundial sobre Contato e Conflito Linguístico, realizado em Bruxelas, inaugurou a vertente (cf. NELDE, 1997 *apud* WINFORD, 2003: 9). Na época, Haugen e Weinreich, particularmente o último, se destacaram ao proporem a abordagem do contato linguístico em sua configuração social (cf. WINFORD, 2003: 9). Aprofundando as investigações, Thomason; Kaufman (1988) se engajaram em resolver antigas controvérsias sobre influências e mecanismos externos e internos ao desenvolvimento de códigos de comunicação. A emergência do método deveu-se, assim, primariamente ao esforço de tais pioneiros (cf. WINFORD, 2003: 9), devido à inexistência de teorias abrangentes sobre o assunto até Weinreich (1953) (cf. CLYNE, 1987: 456 *apud* WINFORD, 2003: 9).

---

<sup>3</sup> Contraindo-se a essas visões, estudos mais recentes (cf. MUFWENE, 2001; WINFORD, 2003, entre outros) disseminam a tese de todas as línguas se submeterem a algum tipo de combinação. Sobre o assunto, Winford (2003: 7) adverte: “os processos de mudança existentes em línguas fortemente mistas tais como crioulos podem ser encontrados em vários níveis nos casos da chamada transmissão ‘normal’. (...) É (...) lamentável que a mudança induzida pelo contato e seus resultados sejam, todavia, vistos por alguns como secundários, e até mesmo marginais, para as buscas centrais da linguística histórico-comparativa”.

<sup>4</sup> O termo “crioulo” advém do nome português “cria”, derivado da palavra latina *creare* (criar). Usado, a princípio, para se referir a um animal domesticado, passou a denotar “escravos nascidos e criados numa colônia da América, por oposição aos nascidos em África” (MOTA, 1996: 525), incluindo mais tarde qualquer mestiço nativo de uma sociedade de base escravocrata e, por conseguinte, suas línguas maternas.

Rejeitando o pensamento tradicional de que o bilinguismo e o convívio entre vernáculos não constituem uma disciplina científica (cf. APPEL; MUYSKEN, 1987: 7 *apud* WINFORD, 2003: 9), Winford (2003: 9) os insere em um campo de estudo específico com temas e objetivos próprios e uma metodologia eclética que envolve abordagens variadas como áreas da sociolinguística e o método histórico-comparativo, por exemplo. De acordo com Weinreich (1953: 5), a interferência entre códigos é estimulada ou bloqueada por fatores estruturais e não estruturais.

Os fatores estruturais são aqueles que se originam na organização de formas linguísticas em um sistema definitivo, diferente para cada língua (...). Os fatores não estruturais são derivados do contato do sistema com o mundo exterior, da familiaridade dos indivíduos com ele, do valor simbólico que (...) é capaz de adquirir e das emoções que pode evocar<sup>5</sup>.

Se múltiplas realidades promovem resultados diversos, cabe à abordagem mais interdisciplinar sustentar e impulsionar a teoria em pauta (cf. WINFORD, 2003: 9). Para a melhor compreensão da natureza e direcionalidade da mudança, devem-se especificar os contextos sociais em que o contato pode ocorrer e examinar o comportamento discursivo dos envolvidos na situação. Trata-se de prever casos de intervenção em uma comunidade bilingue ou multilíngue, definir a estrutura dos falares coexistentes (cf. WEINREICH, 1953: 86) e desvendar as motivações, linguísticas e sociais, das consequências do contato entre falantes de línguas distintas (cf. WINFORD, 2003: 10-11). As próximas seções focalizam o processo de crioulogização, entre outros possíveis produtos do convívio interétnico.

### **3. CRIOLOGÊNESE: PANORAMA SOCIOLINGÜÍSTICO**

Estudos sugerem a relação entre a colonização de terras, o multilinguismo e a crioulogênese. Na era colonial, o convívio entre colonizadores e subjugados e a diversidade linguístico-cultural em um mesmo território compunham uma realidade heterogênea prejudicial para a eleição de um falar oficial. Ao idioma dos europeus (variedade de superstrato ou dominante) costumava-se atribuir um maior prestígio. Representando um alvo de difícil acesso, ele coexistia com as línguas nativas mutuamente ininteligíveis dos povos dominados (variedades mar-

---

<sup>5</sup> Tradução de “The structural factors are those which stem from the organization of linguistic forms into a definite system, different for every language (...). The non-structural factors are derived from the contact of the system with the outer world, from given individuals’ familiarity with the system, and from the symbolic value which (...) is capable of acquiring and the emotions it can evoke”.

ginalizadas ou de substrato), e da necessidade de comunicação imediata surgia um vernáculo, que migrava de segunda língua para primeira, ao se efetivar como língua materna das gerações descendentes (cf. BANDEIRA, 2017: 103).

A ausência de um código comum levava os envolvidos a realizarem adaptações entre os falares coexistentes e delas teriam derivado os chamados *pidgins*. Segundo Winford (2003: 268), eles surgiam

para facilitar a comunicação entre grupos de diferentes famílias linguísticas em contextos restritos tais como comércio, trabalho forçado, e outros tipos de contato marginal. (...) Por definição, (...) são criações adultas, que envolvem (...) aprendizagem e adaptação seletiva dos recursos (...) reminiscetes àqueles encontrados na aquisição de uma segunda língua por adultos.<sup>6</sup>

Representam, então, “vernáculos de contato caracterizados por vocabulário e estrutura fortemente reduzidos, que são nativos para uns, e servem como língua franca”<sup>7</sup> em determinadas situações (WINFORD, 2003: 270). Constituem-se de um inventário lexical repleto de itens genéricos, frequentemente ambíguos e polissêmicos. A expansão do léxico pode, ainda, resultar da “composição, extensão metafórica ou simples empréstimo de palavras da L1 dos falantes ou de uma língua estrangeira”<sup>8</sup> (DRESCHEL’S, 1996 *apud* WINFORD, 2003: 276). Já a fonologia apresenta um conjunto limitado de unidades distintivas, de restrições e de processos, aparentemente resultantes da eliminação de fones e de fonemas não compartilhados pelas línguas em contato, sobretudo daqueles restritos à língua lexicadora principal, que se distanciam da L1 dos aprendizes (cf. WINFORD, 2003: 277-278).

Na crioulistica, é comum relacionar *pidgins* a crioulos, interpretando estes como formas complexas e nativizadas daqueles. Muitos autores defendem a difusão de um *pidgin* como língua materna de uma comunidade e sua evolução, especialmente no discurso de crianças que passam a adotá-lo como primeira língua<sup>9</sup>.

<sup>6</sup> Tradução de “to facilitate communication between groups of different linguistic backgrounds in restricted contexts such as trade, forced labor, and other kinds of marginal contact. (...) By definition, (...) are adult creations, involving processes of learning and selective adaptation of (...) resources (...) reminiscent of those found in adult SLA”.

<sup>7</sup> Tradução de “contact vernaculars characterized by highly reduced vocabulary and structure, which are native to no one, and serve as lingua franca”.

<sup>8</sup> Tradução de “compounding, metaphorical extension or simple borrowing of words from the speaker’s L1 or a foreign language”.

<sup>9</sup> Partindo do pressuposto de que diferentes realidades promovem resultados diversos, Bickerton (1984 *apud* HOLM, 2004: 10) propõe um índice de pidginização, a fim de mensurar a proximidade e o distanciamento das novas línguas de suas fontes. Para o autor, os níveis de afasta-

Por seu turno, outros acrescentam a produção de versões do superstrato como segunda língua, nos estágios iniciais do contato (cf. MUFWENE, 1996a, 1996b; CHAUDENSON, 1989, 1995 *apud* WINFORD, 2003: 332). Os crioulos resultariam da tentativa de adquirir aquele sistema, em meio à escassez de *input* nativo, por razões diversas. A princípio, elaborar-se-ia uma variedade elementar da língua alvo, pautada no processamento do *input* disponível (cf. WINFORD, 2003: 333). Com o acesso cada vez mais limitado às formas originais de referência, o vernáculo passaria a integrar elementos das línguas nativas dos aprendizes, submetendo-se a estratégias compensatórias de processamento (cf. WINFORD, 2003: 334).

A elaboração da gramática crioula envolve um processo complexo de reestruturação, em que os criadores partem da variedade lexificadora, alterando-a ao longo do tempo (cf. WINFORD, 2003: 333). Costuma-se explicar a reestruturação em termos de “relexificação/reanálise” (cf. LEFEBVRE, 1986; LUMSDEN, 1999) ou de “transferência” (cf. SIEGEL, 2003; WEKKER, 1996). Para Winford (2003: 345), ambas se referem a uma reestruturação psicolinguística. No entanto, a

transferência vê os efeitos da influência da L1 da perspectiva da (versões dos aprendizes da) TL [língua alvo], focalizando as maneiras como o *input* (...) é mudado sob essa influência. O cenário da relexificação vê o processo mais da perspectiva da L1 (*input*), focalizando como os itens da L2 são incorporados no sistema dos aprendizes como rótulos para categorias semânticas/funcionais derivadas da L1.(...) Não são em princípio diferentes da retenção das formas lexicais (e às vezes gramaticais) óbvias da L1, que ocorrem tanto na LSA [aquisição de uma segunda língua] (...) quanto na formação de crioulos<sup>10</sup>.

A interferência de fatores externos (sociais, políticos, econômicos e culturais) e internos (linguísticos) no processamento dos dados da língua-alvo, disponibilizados aos aprendizes e utilizados por eles na construção de sua interlíngua, é inerente às elaborações de *pidgins* e de crioulos e à aquisição de uma

---

mento entre os vernáculos emergentes e suas origens podem ser evidenciados tanto no começo quanto no fim do processo de reestruturação.

<sup>10</sup> Tradução de “Transfer views the effects of L1 influence from the perspective of (learner versions of) the TL, focusing on the ways in which input (...) is changed under that influence. The relexification scenario views the process more from the perspective of the L1 (*input*), focusing on how L2 items are incorporated into the learner system as labels for L1-derived semantic/functional categories. (...) are in principle no different from the retention of overt L1 lexical (and sometimes grammatical) forms, which occurs in both LSA (...) and creole formation”.

segunda língua qualquer. Em ambos os casos, desenvolvem-se estruturas funcionais, a partir de informações extraídas da L1 e da L2, o que pode acarretar inovações singulares. Conforme defende Buren (1996: 190), o processo fundamenta-se na sobreposição de uma gramática antiga por novas: o contato com dados relevantes gera a necessidade de acomodar o novo *input*, de que deriva a reestruturação da gramática atual. Na crioulogênese, as mudanças internamente motivadas parecem reproduzir reestruturações de *inputs*, tanto do substrato quanto do superstrato, ambas submetidas a princípios universais de aquisição (cf. WINFORD, 2003: 338). As restrições na habilidade de os indivíduos processarem o *input* da L2 levariam a estratégias de redução e de simplificação. Ressalta-se, entretanto, que, independentemente da variedade a que estejam expostos, todos os recursos das L1s envolvidas na criação de línguas crioulas lhes são disponibilizados, caso precisem compensar o domínio limitado da língua alvo (cf. WINFORD, 2003: 341). Partindo desses pressupostos, a gramática crioula tem sido descrita como reflexo dos parâmetros menos marcados do *input*, da “facilidade de percepção” (“*ease of perception*”), da “facilidade de aprendizagem” (“*ease of learning*”) ou da necessidade de transparência semântica. São hipóteses que explicam alguns de seus princípios, a saber: o isomorfismo entre forma e função (uniformidade); o máximo emprego de regras menos específicas (universalidade); e o mínimo processamento para garantir, através da semântica, estruturas superficiais ou vice-versa (simplicidade) (cf. WINFORD, 2003: 347).

Como os crioulos derivam da associação, em diferentes graus, de características da língua dos colonos e da(s) língua(s) do(s) subalterno(s), constatam-se comportamentos diferentes em uma mesma colônia, ao longo de sua diacronia, e entre colônias de natureza plural, em uma mesma sincronia. Segundo Mintz (1971 *apud* WINFORD, 2003: 312), esses territórios se distinguem quanto: i) ao controle exercido pelas potências europeias sobre a população subalterna; ii) aos códigos usados pelos escravos; e iii) às ideologias das classes dominantes. Em contrapartida, verificam-se similaridades entre as línguas ali faladas, sobretudo, por resultarem de “influências do superstrato e/ou substrato”; de estratégias de “simplificação e gramaticalização”; e da “difusão” ou “migração de falantes crioulos de uma área a outra”<sup>11</sup> (WINFORD, 2003: 308). As múltiplas realidades de que se originam geram peculiaridades entre as línguas de contato e dificultam o estabelecimento de tendências, mas a comparação entre os crioulos evidencia convergências de ordem estrutural (cf. seção 3.1) e sócio-histórica (cf. seção 3.2).

---

<sup>11</sup> Tradução de “superstrate and/or substrate influence, (...) simplification and grammaticalization (...) diffusion, (...) creoles speakers migrated from one area to another”.

### 3.1 As línguas crioulas: tendências linguísticas

À semelhança dos *pidgins*, nos crioulos, o léxico costuma derivar em grande parte do vocabulário da língua dominante (cf. WINFORD, 2003). Certas palavras selecionadas do superstrato passam a representar categorias morfológicas e sintáticas, após se submeterem a processos de gramaticalização ou gramaticização, nos termos de Winford (2003: 350)<sup>12</sup>. Ao mesmo tempo,

alguns itens (...) têm suas formas fonológicas e significados bastante preservados (...). A maioria deles vem de dialetos regionais introduzidos por colonizadores e servos no período inicial de colonização. Mas um número significativo (...) vem da chamada língua ‘náutica’, usada pelos marinheiros e outros navegantes<sup>13</sup> (WINFORD, 2003: 321).

Atesta-se, também, a influência do substrato no inventário lexical, em reanálises semânticas e categóricas das formas originais e na presença de léxico oriundo das línguas dos subalternos. Dessa maneira,

a riqueza dos padrões derivacionais dos crioulos contradiz o mito (...) da ausência de limites morfológicos, uma ideia que sem dúvida surgiu devido à escassez geral de morfologia inflexional (...) resultado de processos de simplificação que crioulos emergentes compartilham com *pidgins* ou (inicialmente) variedades de segunda língua”<sup>14</sup> (WINFORD, 2003: 322).

A comparação entre alguns códigos revela certa variação na incidência das estratégias de simplificação. Por conseguinte, ao mesmo tempo em que, na maioria dos falares, os sistemas pronominais tendem a preservar apenas a distinção de número, sendo desprovidos de caso e de gênero, há variedades mais complexas,

<sup>12</sup> Para o autor, a gramaticalização envolve mudanças graduais internamente motivadas e inclui, de fato, os resultados de mecanismos como a mudança sonora, a reanálise ou analogia. Partindo do pressuposto de que, nas línguas crioulas, esses fenômenos não são graduais, resultando de mudanças relativamente rápidas promovidas “por influências externas (usualmente do substrato), (...) opostas a motivações puramente internas à linguagem”, ele os classifica como fenômenos de gramaticização.

<sup>13</sup> Tradução de “many (...) items have their phonological shapes and meanings fairly well preserved (...). Most of these come from the regional dialects introduced by settlers and servants in the early period of colonization. But a significant number (...) came from so-called ‘nautical’ language, used by sailors and other seafarers”.

<sup>14</sup> Tradução de “The richness of creole derivational patterns belies the myth (...) lack bound morphology, an idea that no doubt arose because (...) generally lack inflectional morphology. (...) result of processes of simplification that emergent creoles share with pidgins, or (early) second language varieties”.

aparentemente por influência do substrato. É o caso do Lung'le e de outras línguas autóctones do Golfo e da Alta da Guiné, que distinguem, por exemplo, no quadro pronominal, formas enfáticas de não enfáticas e acentuadas das átonas (cf. HOLM, 1988: 202 *apud* WINFORD, 2003: 323). No que concerne à morfossintaxe,

talvez a área mais debatida e melhor conhecida (...) dos crioulos seja seu sistema TMA. Com muito poucas exceções, os crioulos expressam significados temporal, modal e aspectual por marcadores pré-verbais. (...) As surpreendentes similaridades levam alguns estudiosos (...) a proporem que um certo inventário bem como uma ordem fixada (T-M-A) de marcadores pré-verbais são definidores dos crioulos. Para tais estudiosos, isso é evidência de que todos os crioulos de fato descendem de um ancestral comum, talvez um *pidgin* português<sup>15</sup> (a respeito da hipótese monogenética, cf. seção 4) (WINFORD, 2003: 324,326)

Grande parte da fonologia dos crioulos parece resultar de reinterpretações, em níveis variados, do superstrato. Os poucos estudos existentes sobre a temática revelam que os falantes tendem a preservar os sons coexistentes nas variedades em contato e, em casos de divergência, pautam-se em regras gerais e nas categorias fonéticas de suas línguas maternas. Frequentemente, incluem na estrutura segmental consoantes plosivas /p t k b d g/ e nasais /m n/. Na existência de diferenças entre elas, os fonemas do superstrato são substituídos por equivalentes encontrados no substrato, o que parece esclarecer, por exemplo, substituições como /b/ por /v/, /t/ por /θ/, /d/ por /ð/, /l/ por /r/ (cf. WINFORD, 2003: 319). No sistema vocálico, “praticamente todos os crioulos têm pelo menos as cinco vogais /i e a o u/, apesar de alguns terem /ε/ e /ɔ/”<sup>16</sup>. São comuns, enfim, “a simplificação de grupos consonantais tanto iniciais quanto finais”; e “o uso da paragoge (a adição de uma vogal especialmente às raízes europeias que terminam em uma consoante)”<sup>17</sup> (WINFORD, 2003: 320).

<sup>15</sup> Tradução de “perhaps the best-known and most debated área of creole (...) is their TMA systems. With very few exceptions, creoles express temporal, modal, and aspectual meanings via pre-verbal markers. (...) The striking similarities led earlier scholars (...) to propose that a certain inventory as well as a fixed (T-M-A) ordering of pre-verbal markers were definitive of creoles. For such scholars, this was evidence that all creoles were in fact descended from a common ancestor, perhaps a Portuguese pidgin”.

<sup>16</sup> Tradução de “practically all creoles have at least the five vowels /i e a o u/, though some have /ε/ and /ɔ/”.

<sup>17</sup> Tradução de “simplification of both initial and final consonant clusters. (...) the use of paragoge (the addition of a vowel especially to European stems that end in a consonant)”.

Talvez no âmbito da sintaxe se verifique a maior distância entre *pidgins* e crioulos. Estes códigos apresentam estruturas mais elaboradas do que aqueles, utilizando “regras de movimento, estratégias de relativização, vários tipos de complementação, e outras formas de subordinação”<sup>18</sup> (WINFORD, 2003: 326). Para Winford (2003: 326), os recursos sintáticos atestados são diferentes dos presentes nos lexicadores. A correspondência entre crioulos mais radicais e seus substratos indica o envolvimento, na constituição da gramática, da “retenção de padrões sintáticos abstratos, aos quais foram incorporadas formas lexicais derivadas do superstrato”<sup>19</sup>. Descartadas certas exceções, eles tendem a preservar muitos elementos da gramática das L1s dos criadores e as divergências encontradas refletem ora essa influência ora desenvolvimentos internos específicos. Em casos particulares como o do crioulo da Guiné-Bissau, a preservação do substrato ao longo do tempo pode estar, parcialmente, relacionada à contínua coexistência da língua emergente com as línguas dos povos dominados (cf. WINFORD, 2003: 336).

*Pidgins* e crioulos “não são versões corruptas ou agramaticais de sua(s) fonte(s) linguística(s), mas sim, línguas legítimas com uma gramática própria”<sup>20</sup> (WINFORD, 2003: 269). As contribuições das variedades dominante e dominadas, respectivamente, na formação do léxico e da gramática dos crioulos, são reconhecidas pela literatura. No entanto, pesquisas têm indicado a frequente coatuação dos dois tipos de *input* nesses processos, revelando que, também, absorvem informações lexicais e estruturais dos substratos, embora essa reprodução não seja exata (cf. WINFORD, 2003: 334). A depender do caso e da interpretação, sua estrutura funcional pode, também, refletir a gramática universal do falante e, em situações peculiares, a língua dominante do território em que surge (cf. seção 4).

Por tais razões, nos últimos anos, tem-se reafirmado a impossibilidade de diferenciar línguas crioulas e não crioulas, com base em suas características estruturais (cf. BANDEIRA, 2017: 107; PRATAS, 2002; BANDEIRA, 2013). Persistindo na tentativa de reuni-las em uma classe linguística identificável, estudiosos partiriam, portanto, de critérios sócio-históricos por natureza (cf. WINFORD, 2003: 308). Como será ressaltado adiante, sob essa perspectiva é a história dos falantes e não a estrutura da linguagem, o primeiro fator determinante para o resultado do contato entre sistemas distintos (cf. HOLM, 2004: 29).

---

<sup>18</sup> Tradução de “movement rules, relativization strategies, various types of complementation, and other forms of subordination”.

<sup>19</sup> Tradução de “the retention of abstract substrate syntactic patterns, into which superstrate-derived lexical forms were incorporated”.

<sup>20</sup> Tradução de “are not corruptions or ungrammatical versions of their source language(s), but rather legitimate languages with a grammar of their own”.

### 3.2 Panorama sociolinguístico da emergência de línguas crioulas: tendências gerais

Do ponto de vista sócio-histórico, os crioulos surgem como códigos nativos de sociedades específicas, como a língua materna de uma população monolíngue ou uma das variedades de falantes multilíngues (cf. BANDEIRA, 2017: 104). São produtos do contato entre indivíduos plurais, linguística, cultural e etnicamente (cf. AFONSO, 2009: 32), que

emergem em configurações coloniais no Novo Mundo, no Oceano Índico e no Oeste da África, onde potências europeias sujeitaram Africanos, Asiáticos e outras populações a suas regras no intervalo entre os séculos XV e XIX. Eles foram criados por escravos e outros grupos subordinados que moldavam materiais das línguas coloniais e suas línguas maternas em um novo meio de comunicação. Eventualmente, (...) vêm a funcionar como vernáculos de uma comunidade adquiridos por crianças como primeiras línguas<sup>21</sup> (WINFORD, 2003: 304).

As trajetórias dos territórios decorrentes da expansão marítima europeia se entrecruzam e indicam a incidência de processos recorrentes, na evolução de novos vernáculos (cf. MUFWENE, 1996: 107). Através de adaptações nas técnicas empregadas pelo método histórico-comparativo à realidade peculiar descrita (cf. DEGRAFF, 2009), eles têm sido definidos e agrupados com base nas coincidências em seu desenvolvimento, a saber: “o isolamento geográfico; o desequilíbrio na demografia; as diferenças na divisão de trabalho e funções entre os cativos (...); a intensidade do contato entre” colonizadores e escravizados; e o papel ativo dos últimos na sociedade (BANDEIRA, 2017: 106). Trata-se de

línguas de contato que emergiram primariamente em áreas de plantação em várias colônias europeias ao redor do mundo. Tais configurações compartilham um número de características sociopolíticas e demográficas, incluindo o uso de um grande número de escravos que foram transportados de suas terras natais e colocados sob o controle de uma pequena minoria de europeus<sup>22</sup> (WINFORD, 2003: 308).

---

<sup>21</sup> Tradução de “emerged in colonial settings in the New World, the Indian Ocean, and West Africa, where European powers subjected African, Asian, and other populations to their rule in the course of the fifteenth through nineteenth centuries. They were created by slaves and other subordinated groups who fashioned materials from the colonial languages and their own mother tongues into new media of communication. Eventually, (...) came to function as community vernaculars acquired by children as first languages”.

<sup>22</sup> Tradução de “contact languages that emerged primarily in plantation settings in various European colonies throughout the world. Such settings shared a number of sociopolitical and demo-

Evoluíram em espaços de exploração da força de trabalho, de dominação política de um grupo sobre outro. Naqueles ambientes, era frequente aniquilarem-se comunidades, submetendo-as a um novo código linguístico e exilando-as nos chamados ambientes-ilha (BANDEIRA, 2017: 102-103). Nas palavras de Araújo (2011: 12), esses espaços correspondem a

uma ilha propriamente dita, uma fazenda do tipo plantation (com o isolamento da população subjugada), um barracão (tanto um galpão na África, onde os escravos chegavam a esperar até dois anos pela chegada dos navios-negreiros que os levariam à América ou os barracões dos entrepostos comerciais, como os encontrados em Curaçao) ou uma fortaleza.

Dos contextos de isolamento supracitados derivam três tipos de crioulos atlânticos: os crioulos de *plantation*, os crioulos de fortaleza e os crioulos de quilombo (cf. BICKERTON, 1988). Os dois primeiros surgiram, respectivamente, em fazendas agroindustriais e em fortalezas ou entrepostos da costa oeste africana, onde os europeus realizavam atividades comerciais. Já os crioulos de quilombo incluem variedades isoladas dos núcleos de colonização, faladas por comunidades de escravos fugidos das plantações (cf. ARENDS, 1995: 16). Para Arends (1995), as semelhanças estruturais entre os vernáculos de *plantation* e de quilombo apontam que aqueles teriam originado estes, embora haja indícios de que o isolamento da comunidade quilombola das colônias tenha afastado as pressões linguísticas europeias e permitido um desenvolvimento mais livre de sua língua.

Interfere, também, na emergência de uma variedade de contato a demografia da região de ascendência. Nos espaços de colonização europeia, era muito destoante o número de colonos e de colonizados, sendo a última camada da população mais representativa (cf. ARENDS, 1995: 19). Por outro lado, quase todos os colonizados eram escravos oriundos de várias regiões, falantes de línguas consideradas inferiores, a quem se atribuíam funções na sociedade. A alguns cabiam atividades exclusivas das plantações (escravos de campo), enquanto outros desempenhavam tarefas domésticas nas casas dos colonos (escravos de casa) e outras práticas como a caça e a pesca. Ademais, poucos cativos atuavam como feitores, preenchendo uma posição intermediária entre brancos e negros. Eles eram responsáveis por castigos, pela imputação de atividades e pelo julgamento da capacidade de doentes preservarem ou retomarem suas funções.

Os cargos exercidos pelos escravos refletiam-se no seu *status* na comunidade e no grau de interação linguística entre eles e os colonizadores (cf. ARENDS,

---

graphic characteristics, including the use of large numbers of slaves who were transplanted from their homelands and placed under the control of a small minority of Europeans”.

1995: 19). Arends (1995: 20) sugere que os fatores ocupariam a posição mais alta de uma hierarquia, sendo sucedidos, respectivamente, pelos escravos de casa, de plantação e improdutivos. Entre a população cativa e os colonos das plantações estariam os chamados “mulatos”, descendentes das relações entre homens brancos e subjugadas. Por vezes, eles eram encaminhados às cidades para atuar nos lares de europeus e, uma vez reconhecidos pelos pais, se tornavam livres. Juntamente com os demais alforriados, a classe representava um estrato intermediário em ascensão, localizado entre a camada restrita de europeus e a grande quantidade de cativos africanos (cf. ARENDS, 1995: 20-21).

Ao longo do tempo, a população subalterna se redistribuiu em boçais (recém-chegados da África), ladinos (veteranos de regiões alheias à colônia) e crioulos (nascidos em cativeiro). Como era comum escravos e não colonizadores transmitirem a língua veicular, a divisão influencia na promoção das línguas de contato (cf. BANDEIRA, 2017: 105). Ao aportarem nas colônias, os boçais traziam consigo uma ou mais línguas africanas. Aos ladinos cabia apresentar-lhes o falar local, enquanto os crioulos o adquiriam como língua materna (geralmente uma língua crioula), mediante o convívio com seus pais e pares. Segundo Arends (1995: 21), em alguns territórios como o Suriname, no começo, a estabilidade das línguas autóctones foi prejudicada pelo intenso fluxo migratório e pelas altas taxas de mortalidade de escravos. Além disso, paulatinamente, crescia o número de cativos nascidos nas colônias e os boçais passaram a ensinar a língua aprendida e utilizada como L2. Esclarece-se, pois, que o novo modelo de aquisição, a segunda língua dos boçais, representava uma variedade muito alterada da língua original.

Os espaços em foco eram marcados por intenso desequilíbrio social e pelo maior prestígio da língua dominante. Sobre o assunto, Winford (2003: 313) alega:

em cada colônia, existiam diferenças no privilégio e status entre as (...) categorias de escravos. Escravos domésticos tinham mais contato com europeus e presumivelmente aprendiam como segunda língua as variedades linguísticas dos últimos. Escravos qualificados tinham mais liberdade de movimento que aqueles que trabalhavam nos campos. Em alguns casos tais diferenças levam a um contínuo linguístico abrangendo desde variedades do superstrato como segunda língua até crioulos altamente divergentes. Similarmente, os crioulos que surgiram em diferentes colônias formam uma espécie de continuum com crioulos mais “radicais” (...) mais afastados dos seus lexicadores, e crioulos intermediários (...) mais próximos<sup>23</sup>.

<sup>23</sup> Tradução de “in each colony, there were differences in privilege and status among (...) categories of slaves. Domestic slaves had more contact with Europeans and presumably learnt second language varieties of the latter’s languages. Skilled slaves had more freedom of movement than

Muitas vezes, nas etapas iniciais da crioulogênese, as variedades autóctones se afastam do superstrato. Ao longo de sua evolução, podem, contudo, caminhar na direção oposta, aproximando-se da língua de referência, em um processo de descrioulização. Constituindo uma das instâncias do contato<sup>24</sup> em que a inserção de uma língua oficial padrão pode ocasionar a perda de traços distintivos, esse fenômeno parece “descrever a distância variável entre a estrutura gramatical de diferentes crioulos e o idioma europeu em que se basearam”<sup>25</sup> (HOLM, 2004: 2). Por sua legítima diversidade, as comunidades crioulófonas podem ser distribuídas, portanto, em um *continuum*<sup>26</sup>, variável em função das línguas lexificadoras. Incluem-se, em seus extremos, os chamados i) basiletos, variedades mais “puras” ou “radicais”, com uma gramática bem distante da língua de superstrato<sup>27</sup>; e ii) acroletos, vernáculos mais semelhantes às línguas dominantes. Por fim, em uma posição intermediária, enquadram-se os iii) mesoletos.

---

those who labored in the field. In many cases, such differences led to linguistic continua ranging from second language varieties of the superstrate to highly divergent creoles. Similarly, the creoles that arose in different colonies form a kind of continuum, with more ‘radical’ creoles (...) furthest removed from their lexifiers, and ‘intermediate’ creoles (...) closer”.

<sup>24</sup> Segundo Holm (2004), paralelamente às línguas crioulizadas e descrioulizadas, existem sistemas entendidos como parcialmente reestruturados. Reconhecem-se abordagens sobre os dois primeiros grupos desde o século XVIII, apesar de apenas no século XX essas pesquisas se intensificarem. O desenvolvimento de um modelo teórico adequado para o tratamento do último caso, promotor de postulações confiáveis sobre a relação entre a história social de seus falantes e a estrutura linguística emergente de um contexto particular é, contudo, mais complexo já que a origem e o desenvolvimento de línguas parcialmente reestruturadas são as principais lacunas dos estudos a respeito do contato linguístico.

<sup>25</sup> Tradução de: “account for the varying distance between the grammatical structure of different creoles and that of the european language they were based” (HOLM, 2004: 2).

<sup>26</sup> A proposta de um *creole continuum* é justificada pelo fato de, em algumas das colônias europeias, o processo de crioulição se estender por várias décadas, enquanto em outras não implicar modificações estruturais relevantes, sendo considerado muito rápido ou até mesmo abrupto. É notório salientar, todavia, que nem sempre é fácil inserir os crioulos nesse *continuum* e identificar “a linha que separa as variedades entre si e do acroleto”. Ademais, é difícil determinar quais características das variedades mais basiletais são inovações motivadas por influências externas e quais refletem a variação inerente à gramática (cf. WINFORD, 2003: 354).

<sup>27</sup> Perspectivas mais tradicionais (cf. MUYSKEN; SMITH, 1995) sugerem que crioulos radicais ou basiletos não se submetem a mudanças regulares, nem tampouco se adequam aos parâmetros utilizados pela linguística histórica, enquanto McWhorter (1998) os insere entre formas prototípicas. Para Mufwene (2001; 67), o estabelecimento de protótipos limita os crioulos “radicais” a um simples grupo dentro das línguas autóctones. Além disso, a ausência de algumas das características tipológicas entre os vernáculos que o integram é recorrente (cf. WINFORD, 2003: 308).

Como se verifica, o contexto é prolixo e estabelecer o(s) principal(is) responsável(is) pela formação e difusão de um novo vernáculo não é uma tarefa simples. A próxima seção focaliza sucintamente o assunto, pautando-se em abordagens superstratistas, substratistas, universalistas e gradualistas.

#### 4. CRIOLOGÊNESE: PRINCIPAIS HIPÓTESES

Em linhas gerais, a criologênese envolve fenômenos tanto individuais quanto coletivos (cf. WINFORD, 2003: 333). O processo é entendido por alguns como reestruturação de uma língua alvo e por outros como complexificação de uma variedade emergencial (cf. seção 2). Seja qual for a interpretação, o objetivo central da crioulística atual tem sido determinar as contribuições das línguas de superstrato e de substrato, além dos universais linguísticos, na formação dos sistemas. A fim de resolver tal impasse, abordagens de cunho superstratista, substratista, universalista e gradualista salientam, respectivamente: i) a língua do colonizador; ii) as línguas dos colonizados; iii) a faculdade da linguagem, inata a todo ser humano e iv) a evolução contínua da gramática crioula (cf. BANDEIRA, 2017: 107).

Considera-se o modelo monogenético o mais extremo da vertente superstratista. Nele, parte-se do pressuposto de que, entre os séculos XV e XVIII, negociantes, contrabandistas e comerciantes utilizavam uma língua de trabalho, cujo léxico era composto majoritariamente por itens do Português (cf. MUYSKEN; SMITH, 1995: 9). Os escravos teriam aprendido o código nos espaços de maior isolamento (ilhas, fortes e navios) e o levavam consigo para as plantações. As línguas crioulas seriam classificadas como falares resultantes de um ancestral comum, provavelmente um *pidgin* português, a que, posteriormente, se acrescentaram elementos de outras línguas europeias. A suposta semelhança entre elas decorreria de sua natureza não complexa, de sua simplicidade e do código subjacente a todas (cf. MUYSKEN; SMITH, 1995: 9-10), o que justificaria, por exemplo, a recorrência de um inventário específico e de uma ordem fixa para tempo, modo e aspecto (TMA) (cf. THOMPSON, 1961; TAYLOR, 1963, 1971).

Trabalhos mais recentes têm, contudo, revelado a falta de evidências linguísticas e históricas para a teoria monogenética (cf. WINFORD, 2003: 326) Mesmo enfatizando a importância das línguas europeias na formação de crioulos, as demais hipóteses superstratistas não se limitam a um único código fundador, acreditando na ocorrência de estratégias de simplificação no discurso destinado a falantes não nativos ou na aprendizagem precária da segunda língua (*baby talk*). Grosso modo, se, na teoria de fala estrangeira, a proximidade entre as línguas autóctones reflete princípios universais regulares na aprendizagem de línguas, na teoria de *baby talk*, a justificativa está no *input* simplificado, isto é, na ocorrência

de adaptações reducionistas no diálogo entre falantes nativos e não nativos. Conforme ressalta Muysken; Smith (1995: 10), seja qual for a diretriz, os superstratistas fundamentam-se na aquisição imperfeita da segunda língua, classificando os crioulos como “estágios fossilizados na seqüência de uma aprendizagem linguística” (BANDEIRA, 2017: 109). Em consonância com Bandeira (2017: 109), defende-se o caráter obsoleto dessas hipóteses, não só por desconsiderarem a quantidade de línguas crioulas do mundo (o que torna a hipótese monogenética infundada), mas também por sustentarem a ideia de uma pseudossimplicidade linguística.

Em oposição aos superstratistas, Mufwene (1996: 87) afirma a influência superior de línguas faladas pelos escravos africanos, na formação dos crioulos. Os novos vernáculos seriam levados para as plantações da região atlântica nos estágios críticos iniciais. Uma das hipóteses da vertente substratista é chamada de “princípio fundador” e se propõe a esclarecer o modo como os falares da primeira população das colônias em que os crioulos surgiram interferem na predeterminação de suas características estruturais (cf. MUFWENE, 2001: 45). Priorizam as marcas dos povos fundadores de uma comunidade, por serem as mais antigas empregadas por um maior contingente de falantes. Ademais, tal herança linguística seria transmitida entre os descendentes, sendo preservada ao longo do tempo (cf. MUFWENE 2001: 28-29; FREITAS, 2016: 169).

O conceito de efeito fundador, postulado pela genética de populações, remete ao desprendimento de um pequeno grupo em um maior: na tentativa de colonizar novos espaços, trazem consigo parte da diversidade existente na camada emergente. Linguisticamente, variedades mais diversas tenderiam a se aproximar da origem, ao passo que aquelas mais semelhantes se distanciariam (cf. BOUCKAERT et al., 2012). Para Mufwene (1996: 84), aos criadores da comunidade de fala, utentes tanto de superestrato quanto de substrato, atribuir-se-ia um papel mais relevante, no que diz respeito à seleção de componentes estruturais ao longo da formação da língua de contato. Compartilhando essa possibilidade, Bandeira (2017: 110) salienta:

na fase de plantação, (...) a parcela recém-chegada, pela demanda do regime escravocrata, buscava aprender a língua local da melhor forma possível, não planejando, por conseguinte, falar conforme os princípios de suas línguas nativas ou quaisquer outras faladas antes (...). Logo, a fala da população fundadora se torna alvo desse contingente. Como consequência, os recursos linguísticos já empregados pela população fundadora tornam-se suscetíveis de serem transmitidos por meio de seleções e adaptações sucessivas, o que define a evolução gradual de novas línguas.

Se a variedade de superestrato aos poucos foi apropriada por falantes de múltiplas origens, há de se reconhecer a influência dos falares desses naquela, “da

mesma maneira como diversas línguas não crioulas foram influenciadas por outras em suas histórias” (BANDEIRA, 2017: 110). Quando as variedades de substrato interferem decisivamente na seleção entre as opções disponíveis na de superstrato, verifica-se um processo de reorganização do código. Por conseguinte, a maioria dos componentes escolhidos para compor o inventário das línguas crioulas se originaria de ambas (cf. MUFWENE, 2001: 232): os crioulos surgiriam de tentativas dos colonizados de falar a língua de superstrato e de adaptações das duas partes envolvidas no cenário de contato (cf. MUFWENE, 1996, 2001). Decorreriam de experiências das populações coexistentes: tentando se comunicar na língua dos colonos não dominada por quase todos, passariam a adotar um meio de comunicação interétnico (cf. MUFWENE, 2001: 51).

Em dissonância com as perspectivas supracitadas, há análises descritivas sobre as estruturas dos crioulos fundamentadas nos mecanismos inatos da faculdade humana da linguagem. Sob esse ponto de vista,

aquelas características da gramática (...) que surgem via processos internos e não tem contraparte (...) nas fontes de substrato e superstrato (...) podem ser devidas tanto à criatividade no próprio processo de formação do crioulo, quanto a mudanças subsequentes internamente motivadas. Em ambos os casos, as crianças podem ter desempenhado um papel vital<sup>28</sup> (WINFORD, 2003: 348).

Dentre outras interpretações da corrente universalista, sobretudo inatista, salienta-se a teoria do Bioprograma, por propor que os crioulos resultariam essencialmente de criações infantis, após a aquisição de um *pidgin* corrompido, desenvolvido por adultos nos estágios iniciais de colonização (cf. BICKERTON, 1981). Elas teriam acesso a um código incompleto, deficiente, transmitido pela população adulta com mínimo ou nulo acesso a falantes nativos da língua de superstrato. O *input* recebido seria desprovido de estruturas suficientes para atuar como língua natural e ser transformado em uma língua completa, através, exclusivamente, das capacidades linguísticas inatas das crianças.

De acordo com Muysken; Smith (1995: 11 *apud* BANDEIRA, 2017: 112), apesar de substratistas admitirem a atuação de alguns dos mecanismos inatos descritos na teoria do Bioprograma, Bickerton rejeita qualquer influência do substrato na formação dos crioulos. Sobre o assunto, Holm (2000: 64) ratifica que a exclusão na íntegra do papel do substrato suscita um maior esforço da cor-

---

<sup>28</sup> Tradução de “those features of creole grammar (...) that arise via internal processes and have no counterparts in (...) substrate or superstrate sources (...) may be due either to creativity in the process of creole formation itself, or to subsequent internally motivated change. In both cases, children may have played a vital role”.

rente substratista para demonstrar a relevância das línguas africanas na emergência dos múltiplos falares distribuídos pelo Atlântico. Por essas e outras razões, atualmente, tem-se refutado o extremismo da hipótese do Bioprograma e defendido que os princípios da gramática universal atuam em todas as etapas do desenvolvimento, se apresentando desde a fase inicial *pidgin* até os últimos estágios. Cabe-lhe, todavia, “restringir o processo de reestruturação pelo qual *inputs* do superstrato e do substrato dão forma a uma gramática viável”<sup>29</sup> (WINFORD, 2003: 347).

Se os superstratistas se pautam em conceitos ultrapassados, substratistas não apresentam dados empíricos para explicar as semelhanças estruturais entre as línguas de contato e universalistas falham ao desconsiderarem qualquer influência social, uma característica inerente a todas as línguas. Propostas gradualistas, por seu turno, preocupam-se em explicar a estabilidade da gramática das línguas crioulas, em meio às variantes em competição, introduzidas por diferentes indivíduos ou grupos de falantes (cf. WINFORD, 2003: 345). Afirmam que o regime de plantação intensificava as demandas comunicativas, pois os escravos recém-chegados não conheciam a língua comum. Daí, resultariam sistemas de comunicação improvisados. Alguns deles se expandiriam, a partir do uso contínuo, e se estabilizariam como crioulos. Trata-se de um processo gradual em que crianças e adultos obedecem a princípios da gramática universal que orientam a construção do novo código. Ao adquirirem a língua alvo como segunda língua, “adultos introduzem padrões inovadores na ecologia (...) das línguas dos aprendizes, enquanto as crianças via aquisição de uma primeira língua, desempenham um importante papel na reestruturação das inovações dos adultos (e suas próprias) em uma gramática estável”<sup>30</sup> (DEGRAFF, 1999b: 494). Para Winford (2003: 346), nos espaços em que os substratos coexistentes convergem, a gramática da língua emergente apresenta muitos traços da L1. Sobre a influência do substrato nos crioulos do Golfo da Guiné, Hagemeijer (2011: 36) adverte: “é previsível que um substrato relativamente homogêneo e cada vez menos o modelo do superstrato, devido à mudança gradual em direção à sociedade de plantação, favoreça a introdução de características do substrato (mais específicas)” (cf. SINGLER, 1988; THOMASON; KAUFMAN, 1988: 157-8).<sup>31</sup> As paridades entre as estruturas se-

<sup>29</sup> Tradução de “to constrain the processes of restructuring by which superstrate and substrate inputs are shaped into a viable grammar”.

<sup>30</sup> Tradução de “adults introduce innovative patterns into the (...) ecology of language learners, whereas children, via L1A (first language acquisition), play a key role in restructuring adults’ (and their own) innovations into stable grammars”.

<sup>31</sup> Tradução de “it is predictable that a relatively homogeneous (founder) substratum and increasingly less access to the superstrate model due to the gradual shift towards a plantation society

riam, portanto, produto de um contexto social comum, e não, de aspectos cognitivos ou das línguas de substrato (cf. MUYSKEN; SMITH, 1995: 11).

## 5. CONSIDERAÇÕES FINAIS

É evidente que alguns aspectos relativos ao convívio entre distintos falares são relevantes para descrever realidades multilíngues. No entanto, por se tratar de contextos plurais e ao mesmo tempo singulares, acredita-se que as línguas autóctones ali faladas não decorreriam, exclusivamente, das variedades de superstrato e de substrato ou de mecanismos linguísticos universais disponíveis a qualquer falante em uma faculdade inata. Na verdade, podem derivar da atuação conjunta de todos os fatores, de elementos sociais e das “histórias dos indivíduos (...) obrigados a aprender um código enquanto (...) cativos, alijados de seus direitos mais básicos” (BANDEIRA, 2017: 113). Os utentes seriam responsáveis pelo desenvolvimento dos vernáculos emergentes em sistemas de comunicação plenos (e não ‘simples’), diferentes daqueles envolvidos no contato. O processo compreende vários *inputs* e estratégias de reestruturação, mas a maioria das teorias sobre a crioulogênese tendem a se limitar a aspectos específicos. Assim, é um desafio para os estudos contemporâneos integrar essas considerações por vezes paradoxais em uma abordagem coerente sobre a origem dos crioulos (cf. WINFORD, 2003: 340).

## REFERÊNCIAS

- AFONSO, H. L. *Interferências linguísticas: um contributo para o ensino da língua portuguesa em S. Tomé e Príncipe*. 2009. 184f. Dissertação (Mestrado em Língua Portuguesa) – Faculdade de Letras, Universidade de Lisboa, Lisboa, 2009.
- ARAUJO, G. A. Empréstimos recentes do Português, variação fonética e a sílaba na língua são-tomense da ilha de São Tomé. *Papia*, São Paulo, v. 17, p. 55-66, 2007.
- ARAUJO, G. A. *Três textos em Papiamentu clássico*. 2011. Tese (livre-docência em Filologia e Língua Portuguesa) – Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo, 2011.
- ARENDS, J. The socio-historical background of creoles. In: ARENDS, J.; MUYSKEN, P.; SMITH, N. (Eds.). *Pidgins and creoles: an introduction*. Amsterdam/Filadélfia: John Benjamins Publishing Company, 1995. p. 15-24.
- BANDEIRA, M. *A adaptação de empréstimos recentes no Papiamentu moderno*. 2013. 2013. 234f. Dissertação (Mestrado em Filologia e Língua Portuguesa) – Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo, 2013.

---

favour the introduction of (more and specific) substrate-induced features (e.g. Singler 1988; Thomason & Kaufman 1988: 157-8).

- BANDEIRA, M. *Reconstrução fonológica e lexical do protocioulo do Golfo da Guiné*. 2017. 437f. Tese (Doutorado em Filologia e Língua Portuguesa) – Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo, 2017.
- BICKERTON, D. *Roots of language*. Ann Arbor: Karoma, 1981.
- BICKERTON, D. Creole languages and the bioprogram. In: NEWMeyer, F. J. (Ed.). *Linguistics: The Cambridge survey*. v. 11; *Linguistic Theory: extensions and implications* Cambridge. Cambridge: Cambridge University Press, 1988. p. 268-284.
- BOUCKAERT, R. *et al.* 2012. Mapping the origins and expansion of the indo-european language family. *Science*, USA, v. 337, p. 957-960, 2012.
- BUREN, P. Are there principles of universal grammar that do not apply to second language acquisition? In: JORDENS, P.; LALLEMAN, J. (Eds.). *Investigating second language acquisition*. Berlin: Mouton de Gruyter, 1996. p. 187-207.
- DEGRAFF, M. Creolization, language change and language acquisition: an epilogue. In: DEGRAFF, M. (Ed.). *Language creation and language change: creolization, diachrony and development*. Cambridge, MA: MIT Press, 1999b. p. 473-543.
- DEGRAFF, M. Language acquisition in creolization and, thus, language change: Some cartesian-uniformitarian boundary conditions. *Language and Linguistics Compass*, New Hampshire, v. 3/4, p. 888-971, 2009.
- FREITAS, S. *Contribuições linguísticas cabo-verdiana e sefardita na formação do Papiamentu*. 2016. 669f. Tese (Doutorado em Filologia e Língua Portuguesa) – Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo, 2016.
- HAGEMEIJER, T. The Gulf of Guinea creoles: genetic and typological relations. *Journal of Pidgin and Creole Languages*, v. 26, n. 1, p. 111-154, 2011.
- HOLM, J. *Pidgins and creoles*. Cambridge: Cambridge University Press, 1988.
- HOLM, J. *An introduction to pidgin and creoles*. Cambridge: Cambridge University Press, 2000.
- HOLM, J. *Languages in contact: the partial restructuring of vernaculars*. Cambridge: Cambridge University Press, 2004.
- LEFEBVRE, C. Relexification in creole genesis revisited: The case of haitian creole. In: MUYSKEN, P.; SMITH, N. (Eds.). *Substrata versus universals in creole genesis*. Amsterdam/Filadélfia: John Benjamins Publishing Company, 1986. p. 279-301.
- LUCCHESI, D.; BAXTER, A.; RIBEIRO, I. (Eds.). *O Português afro-brasileiro*. Salvador: EDUFBA, 2009.
- LUMSDEN, J. S. The role of relexification in creole genesis. *Journal of pidgin and creole languages*, v. 14, p. 225-258, 1999.
- MCWHORTER, J. Identifying the creole prototype: vindicating a typological class. *Language*, Philadelphia: University of Pennsylvania, v. 74, n. 4, p. 788-818, 1998.
- MEILLET, A. *Linguistique historique et linguistique générale*. Paris: Champion, 1921. v. 1.

MOTA, M. A. C. da. Línguas em contacto e variação. In: FARIA *et al.* *Introdução à linguística geral e portuguesa*. 2. ed. Lisboa: Editorial Caminho, 1996.

MOTA, J. A.; CARDOSO, S. A. M. Para uma nova divisão dos estudos dialetais brasileiros. In: MOTA, J. A.; CARDOSO, S. A. M. (Eds.). *Documentos 2: Projeto Atlas Linguístico do Brasil*. Salvador: Quarteto, 2006. p. 15-34.

MUFWENE, S. The development of American Englishes: some questions from a creole genesis perspective. In: SCHNEIDER, E. W. (Ed.). *Focus on the USA [varieties of English around the world]*. Amsterdam: John Benjamins, 1996a. p. 231-264.

MUFWENE, S. The founder principle in creole genesis. *Diachronica*, v. 13, p. 83-134, 1996b.

MUFWENE, S. The founder principle in the development of creoles. In: MUFWENE, S. *The ecology of language evolution*. Cambridge: Cambridge University Press, 2001. p. 25-80.

MÜLLER, M. *Lectures on the science of language*. London: Longman, Green Longman and Roberts, 1861.

MUYSKEN, P.; SMITH, N. The study of pidgin and creole languages. In: ARENDS, J.; MUYSKEN, P.; SMITH, N. (Eds.). *Pidgins and creoles: an introduction*. Amsterdam/Filadélfia: John Benjamins Publishing Company, 1995. p. 3-14.

NASCIMENTO, Fabiane de Mello Vianna da Rocha Teixeira Rodrigues. *O sistema vocálico do Português de São Tomé e o comportamento das vogais médias em contexto pretônico*. 2018.194f. Tese (Doutorado em Letras Vernáculas). Universidade Federal do Rio de Janeiro, Faculdade de Letras, Rio de Janeiro, 2018.

OKSAAR, E. Bilingualism. In: SEBEEK, T. A. (Ed.). *Current trends in linguistics*. v. 9; Linguistics in Western Europe. Den Haag: Hague University, 1972a. p. 476-511.

ORLANDI, E. P. Processo de descolonização linguística: as representações da língua nacional. In: GALVES, C.; GARMES, H.; RIBEIRO, F. R. (Eds.). *África-Brasil: caminhos da língua portuguesa*. Campinas: Editora da Unicamp, 2014.

PRATAS, F. *O sistema pronominal do caboverdiano (variante de Santiago)*. 2002. 129f. Dissertação (Mestrado em Linguística) – Faculdade de Ciências Sociais e Humanas, Universidade Nova de Lisboa, Lisboa, 2002.

SCHMIDT, J. *Die Verwandtschaftsverhältnisse der Indogermanischen Sprachen*. Weimar: Hermann Böhlau, 1872.

SCHUCHARDT, H. *Ueber des Negerportugiesische von S. Thomé*. Sitzungsberichte Wien, v. 101, n. 2, 1882. p. 889-917.

SINGLER, J. The homogeneity of the substrate as a factor in pidgin/creole genesis. *Language*, Philadelphia: University of Pennsylvania, v. 64, p. 27-51, 1988.

TAYLOR, D. The origin of west indian creole languages: evidence from grammatical categories. *American Anthropologist*, Denver, v. 65, p. 800-814, 1963.

TAYLOR, D. Grammatical and lexical affinities of creoles. In: HYMES, D. *Pidgnization and creolization of languages*. Guent: Story, 1971. p. 293-296.

THOMASON, S. G.; KAUFMAN, T. *Language contact, creolization, and genetic linguistics*. Berkeley: University of California Press, 1988.

THOMPSON, R. W. A note on some possible affinities between the creole dialects of the Old World and those of the New. In: POBERT, B. L. *LePage creole language studies: proceedings of the conference on creole language studies*. London: MacMillan, 1961. p. 107-113.

WEINREICH, U. *Languages in contact, findings and problems*. New York: Mouton, 1953.

WEKKER, H. Creolization and the acquisition of English as a second language. In: WEKKER, H. (Ed.). *Creole languages and language acquisition*. Berlin: Mouton de Gruyter, 1996. p. 139-149.

WINFORD, D. *An introduction to contact linguistics*. Oxford: Blackwell Publishing, 2003.

# 2

## CAPÍTULO

# SÃO TOMÉ E PRÍNCIPE: aspectos históricos, econômico-sociais e linguísticos

**FABIANE DE MELLO VIANNA DA ROCHA  
TEIXEIRA RODRIGUES DO NASCIMENTO**

## **1. INTRODUÇÃO**

Este capítulo<sup>1</sup> expõe algumas informações históricas, socioeconômicas e linguísticas a respeito das ilhas de São Tomé e Príncipe, na tentativa de traçar um perfil sociolinguístico da região e explicar sua evolução desde o descobrimento até os dias atuais. Como ilustram as Figuras 1 e 2, São Tomé associa-se a Príncipe em um estado insular situado a aproximadamente 250 Km da costa africana, no Golfo da Guiné.

---

<sup>1</sup> Este capítulo foi integralmente reproduzido da tese de Nascimento (2018).



**Figura 1** Localização de São Tomé, Príncipe e Ano Bom no Golfo da Guiné<sup>2</sup>



**Figura 2** Distritos de São Tomé e de Príncipe<sup>3</sup>

<sup>2</sup> Imagem extraída de <https://pt.wikipedia.org/wiki/Ano-Bom>.

<sup>3</sup> Imagem extraída de [https://www.google.pt/search?q=distritos+caue&source=lnms&tbn=isch&sa=X&ved=0ahUKEwiFx7Li9LzUAhXK1hoKHUf4DDsQ\\_AUIBigB&biw=1371&bih=571](https://www.google.pt/search?q=distritos+caue&source=lnms&tbn=isch&sa=X&ved=0ahUKEwiFx7Li9LzUAhXK1hoKHUf4DDsQ_AUIBigB&biw=1371&bih=571)

A escassez de documentos históricos dificulta a determinação das descobertas, mas estudos sobre o assunto remetem ao fim do século XV, relacionando suas trajetórias à de Ano Bom.<sup>4</sup> 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 povoamentos 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). 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) (cf. seção 2.2).

Segundo Ladhams (2007 *apud* BANDEIRA, 2017: 120), “o input linguístico apresentado pelo influxo cativo nos primeiros anos de colonização de São Tomé é complexo e possui contornos de difícil precisão”. Desde o fim do século XV, a necessidade de diálogo entre colonizador e colonizados de diferentes origens promoveu um falar emergencial, que, uma vez complexificado e nativizado, resultou no PCGG. Representando frutos da especiação e expansão da língua-mãe, além do Português, do Caboverdiano e de resquícios de línguas bantas, coexistem em São Tomé i) o Forro (ou santomé), crioulo mais usado na ilha, e ii) o Angolar, utilizado no sudoeste e nos distritos de Cauê e Lembá, principalmente. Nas proximidades, encontram-se utentes iii) do Lung’ie, em Príncipe; e iv) do Fa d’ambô, cuja comunidade linguística se concentra em Ano Bom, a cerca de 150 km do Sul do país, no Oeste da África<sup>5</sup> (cf. FERRAZ, 1979; HAGEMEIJER, 2009, 2001; GONÇALVES, 2016) (cf. seção 3.1).

De acordo com Hagemeijer (2009: 15), “a breve incursão na sintaxe não só mostra que os crioulos do Golfo da Guiné têm essencialmente a mesma estrutura mas também que as estratégias sintáticas se inclinam, em geral, para estratégias

---

&dpr=1.4#tbn=isch&q=distritos+sao+tome+e+principe+caue&imgrc=96ytzIKYZKe9XM:&spf=1497428761804

<sup>4</sup> De acordo com Ferraz (1979: 9), o espaço era desabitado “até a chegada dos portugueses em datas que são geralmente situadas entre 1470 e 1472.” Contrariando essa visão, Campos (1971 *apud* Ferraz, 1979: 9) afirma que “São Tomé foi descoberta em 1478, Príncipe em 1479 e Ano Bom oito anos depois”.

<sup>5</sup> Atualmente, também se encontram falantes de Fa d’ambô na Ilha de Bioko e na diáspora (cf. HAGEMEIJER; ZAMORA; 2016: 193).

(...) disponíveis no substrato Edo”<sup>6</sup>. Trata-se de sistemas que aliam um léxico predominantemente português (cerca de 90%) a estruturas funcionais incorporadas da gramática das línguas africanas de que se formaram (GONÇALVES, 2016: 24). A gênese singular contrasta com a inteligibilidade limitada entre eles, fato justificável pela “separação no tempo e no espaço” e “pelas mudanças internas e externas (como contato linguístico posterior) que cada um sofreu depois de se ter separado do tronco comum” (HAGEMEIJER, no prelo a). Por isso, sobretudo nos níveis lexical e fonológico, o angolar, “apresentaria componentes mais marcadamente banto, ao passo que o Lung’ie, devido ao Príncipe ter recebido um maior número de escravos do Delta do Níger, apresentaria um componente mais marcadamente edo” (BANDEIRA, 2017: 119). Ao mesmo tempo, o Fa d’ambô compartilha 82% de seu léxico com o Forro. A influência africana nesses falares, “inferior a 10% do léxico global, é primeiramente derivada do Edo, a linguagem do antigo Reino de Benin na Nigéria, e do Quicongo, um grupo de línguas do oeste Banto falado no Congo e no Norte da Angola”<sup>7</sup> (HAGEMEIJER; ZAMORA, 2016: 197).

As próximas seções aprofundam as informações históricas e linguísticas listadas, concentrando-se (seção 2) na fase de habitação e nos ciclos do açúcar (2.1), do café e do cacau (seção 2.2). No que tange ao multilinguismo (seção 3), expõem-se dados a respeito dos crioulos do Golfo da Guiné (3.1) e da(s) variedade(s) do Português encontrada(s) no país (cf. seção 3.2). Na seção 4 tecem-se as considerações finais.

## 2. CONTEXTUALIZAÇÃO HISTÓRICA

### 2.1 A primeira colonização e o ciclo do açúcar

É consensual entre as pesquisas inseridas na bibliografia deste estudo a afirmativa de que São Tomé, “o núcleo das três ilhas, foi a primeira a ser colonizada”

<sup>6</sup> Em consonância com Ladhams (2007), Bandeira (2017: 117) sugere que os primeiros escravos “nem sempre eram edo e, por conseguinte, não falavam necessariamente” línguas edoídes. Sob o domínio do Reino de Benin, encontravam-se grupos etnolinguísticos diferentes: haveria à época uma região ocupada pelo itsekiri, falantes de uma língua ioruboide, em conjunto com o porto de Ughoton e o Rio de Forcados. Ademais, “os mercadores de escravos podem ter sequestrado mão-de-obra fora da região linguística do Delta do Níger” (BANDEIRA, 2017: 117). Sendo assim, embora os dados linguísticos fortaleçam a hipótese de predomínio edoide, o ijoide também não deve ser descartado (cf. HAGEMEIJER, 2011).

<sup>7</sup> Tradução de “at less than 10 percent of the overall lexicon, is primarily derived from Edo, the language of the old kingdom of Benin in Nigeria, and Kikongo, a western Banto language cluster spoken in the Congo and northern Angola”.

(FERRAZ, 1979: 9). A ilha estaria desabitada, quando ocorreu uma tentativa fracassada de povoá-la em 1486. Outras expedições aconteceram<sup>8</sup>, mas apenas em 1493, um grande contingente de colonos migra de Portugal. Trata-se do ano da nomeação de Álvaro de Caminha como capitão-donatário, cargo mantido até 1505. Logo, encontrar-se-iam, ali, além de uma maioria de escravos, brancos a serviço da coroa portuguesa ou do comércio com a região da Guiné (BRÁSIO, 1953 *apud* BANDEIRA, 2017: 115).

Dados históricos e linguísticos comprovam relações entre Portugal e o Reino de Benin, entre os séculos XV e XVI, indicando que “a primeira colonização em São Tomé recorreu predominantemente, se não exclusivamente, à importação de escravos do Delta do Níger”<sup>9</sup> (atual Nigéria). Associaram-se a eles muitos degredados encaminhados para lá, após cometerem crimes em Portugal. Do mesmo país, teriam vindo, ainda, em 1493, 2.000 crianças judias para se converterem ao catolicismo, depois de os pais serem expulsos da Espanha (cf. FERRAZ, 1979: 16; HAGEMEIJER, 2011 a: 113; BANDEIRA, 2017: 115). No entanto, grande parte faleceu pela insalubridade das embarcações ou por doenças adquiridas na chegada, restando, em 1506, cerca de 600 meninos e meninas<sup>10</sup> (cf. BRÁSIO, 1953: 33-45 *apud* BANDEIRA, 2017: 115).

A primeira colonização parece ter estreitado o relacionamento entre europeus e africanos, se comparadas as fases da trajetória sociolinguística da ilha. A convivência mais intensa entre portugueses e escravos com suas respectivas línguas maternas (edoides e bantas, dentre as quais se destacam variedades do Quicongo e Quimbundo), no fim do século XV e nos primórdios do XVI, teria implicado uma aproximação dos subalternos ao código linguístico utilizado pelos

---

<sup>8</sup> De acordo com Caldeira (2008: 50), em meio a doenças contagiosas comuns na Europa como a cólera e a febre tifoide, “a malária (...) pelo seu efeito dizimador, criará a ideia das ilhas de São Tomé e Príncipe como “cemitério de europeus”, afastando potenciais candidatos à fixação no arquipélago e, correlativamente, facilitando a africanização do território. De facto, as populações que crescem nas zonas endémicas (...) desenvolvem anticorpos que lhes permitem ganhar uma relativa imunidade (...), o que de todo não acontecia com europeus acabados de desembarcar”.

<sup>9</sup> Tradução de “the early settlement on S. Tome relied predominantly, if not exclusively, on slave imports from the Niger Delta”.

<sup>10</sup> Pautado em um manuscrito de Valentim Fernandes, Sousa (1990 *apud* HAGEMEIJER, 2011) ressalta que havia, em 1506, na ilha de São Tomé, 1.000 moradores de residência fixa (europeus principalmente portugueses), 2.000 escravos permanentes e cerca de 5.000 ou 6.000 escravos transitórios. Entre os primeiros encontram-se, uma maioria de degredados, mas também membros do clero, soldados, outros homens, mulheres livres, e possivelmente também um grande número de órfãos judeus de Seville, que deviam ter menos de 8 anos de idade quando chegaram em 1493.

colonos (cf. HAGEMEIJER, 2009: 2). Complementando essa perspectiva, Clements (2000: 187) sugere a existência entre os últimos de “uma consciência alargada de *foreigner talk*” do idioma, ao alegar:

a gente daquela época conhecia estratégias de simplificação da sua própria língua, conhecimento que vinha dos portugueses e a grande quantidade de estrangeiros que tinham vindo para Portugal para repovoar a terra conquistada aos mouros pelo exercício cristão.(...) Nas cidades (...) da costa (Lisboa, Porto) existiam provavelmente uma ou mais variedades simplificadas do português que se deveriam ter formado e utilizado entre marinheiros e comerciantes. Essas variedades coexistiam com outras variedades linguísticas que se empregavam na comunicação, como por exemplo a língua franca.

Da necessidade iminente de diálogo entre colonos, falantes da língua portuguesa, e colonizados, em geral utentes de línguas africanas, teria resultado uma variedade *pidgin* e as diversas estratégias de simplificação certamente interferiram na sua formação (CLEMENTS, 2000: 187).

Acrescem-se a esses aspectos, as políticas de povoamento. Segundo Cortesão (1968: 33), havia desde o princípio, um decreto estabelecendo a doação de uma mulher escrava para cada condenado exilado. Anos depois, o rei Manuel determina a liberdade delas e dos descendentes e surgem os chamados “filhos da terra” ou “forros”, uma classe composta por moradores de São Tomé, alguns, inclusive, mais abastados e poderosos, detentores de muitos subordinados (ANÔNIMO, 1550: 52; *apud* BANDEIRA, 2017: 118). Para Hagemeijer (2009: 4), é plausível a participação dessa comunidade “na origem e consolidação da nova língua que se falava na ilha.

A ausência de “uma atividade econômica de envergadura” e poucos povoadores tornavam a mão-de-obra essencialmente africana nas mais variadas tarefas (rurais, domésticas e de obras – cf. HAGEMEIJER, 2009: 2). Dividiam-se, pois, os cativos em dois estratos principais: os escravos de casa, com morada permanente na ilha; e os escravos de resgate, sequestrados de regiões do litoral africano e vendidos como mercadoria (cf. HAGEMEIJER, 1999; HLIBOWICKA-WEGLARZ, 2012; BANDEIRA, 2017). Aparentemente, o *status* local daqueles também interferiria na formação da língua emergente e boa parte dos demais<sup>11</sup> seria reexportada para a feitoria da Mina como moedas de troca. A falta de embarcações estendia, contudo, a estada dos resgatados para além dos cinquenta dias es-

---

<sup>11</sup> De acordo com Hlibowicka-Weglarz (2012: 178 *apud* BANDEIRA, 2017: 116), apenas uma pequena parcela dos escravos de resgate permanecia na ilha para pagar os soldos aos primeiros colonizadores.

timados. Por isso, “passavam muitos meses em São Tomé, trabalhando como mão-de-obra temporária nas plantações e, participando, de certo modo, do processo de criouliização” (BANDEIRA, 2017: 116).

A partir de 1500, os portugueses garantem direitos sobre os cativos. O poderio exclusivo só foi atingido, contudo, em 1515, quando São Tomé se tornou um importante entreposto de escravos destinados a trabalhar nas plantações de açúcar e a “força motriz” (“*driving force*”) das ilhas do Golfo da Guiné (cf. HAGEMEIJER; ZAMORA, 2016: 195).

Há relatos sobre plantações em São Tomé, desde os decretos de 1485, embora o ciclo econômico de cultivo do açúcar tenha se iniciado somente na década de 1510. Estendeu-se ao longo do século XVI, com ápice nos anos 1560, após a ilha ser reconhecida como exportadora, e declinou na primeira metade do século XVII, quando a maioria dos plantadores parte para o Brasil (cf. HAGEMEIJER, 2011 a; HAGEMEIJER; ZAMORA, 2016; HAGEMEIJER, no prelo). Aparentemente, aspectos como o “clima quente com chuvas abundantes, solo vulcânico, por conseguinte, fértil e (...) a disponibilidade do emprego de uma larga escala de cativos” (BANDEIRA, 2017: 119) favoreciam o plantio na região. As mesmas características devem, também, ter interferido no fracasso, pois, em relação ao fabrico, o excesso de humidade prejudicava “a secagem dos ‘blocos de açúcar’”. À qualidade superior do produto no novo mundo, somaram-se fatores como i) “a ameaça representada por outras nações europeias que começaram a navegar pelo Golfo da Guiné”; iii) “ataques ao sistema de plantação santomense por escravos fugidos”<sup>12</sup> (HAGEMEIJER, no prelo); além iv) da “paritose que afectou as plantações (...) entre 1580 a 1595” (CALDEIRA, 2008: 50).

Durante a primeira colonização, a língua emergencial se complexificou e se efetivou como nativa, dando origem ao chamado PCGG: o Santomé ou Forro original, crioulo de base lexical portuguesa dominante na ilha até os dias atuais (HAGEMEIJER, no prelo). No mesmo período, parte da população migrou para povoar o Príncipe, estendendo-se a Ano Bom, na metade do século XVI. A necessidade de trabalhadores nas plantações sacarinas intensificou a exploração de mão-obra escrava. Congo e Angola se destacaram como áreas de resgate, e, rapidamente, declinou o poderio da Nigéria. Se “o predomínio de escravos Edo coincide com (...) o período de povoamento, que durou de 1485 até aproximadamente 1510/1520”<sup>13</sup>, os escravos Banto chegam em uma fase posterior à cristalização do

---

<sup>12</sup> Tradução de “the threat posed by other European nations who started navigating the Gulf of Guinea, and the attacks on the Santomean plantation system by runaway slaves”.

<sup>13</sup> Tradução de “The predominance of Edo slaves overlaps (...) the homestead period, which lasted from 1485 until approximately 1510/1520”.

PCGG (HAGEMEIJER, 2011: 112). Mais precisamente, “a mudança para uma sociedade de plantação corresponde à emergência e subseqüente domínio de áreas Banto, primeiro do Congo e depois da Angola, no tráfico de escravos para São Tomé”<sup>14</sup> (HAGEMEIJER, 2011: 112). Os novos cativos falavam, especialmente, línguas como Quicongo e o Quimbundo e aumentam em número nos primeiros anos da fase de plantação (anos 1510) com um impacto permanente:

Estudos genéticos realizados em São Tomé e Príncipe apresentam a mistura das populações que se tornaram os habitantes das ilhas ao longo dos séculos. Enquanto o genótipo europeu representa apenas 10% do genótipo santomense, o genótipo africano constitui os outros 90% e pode ser mais especificado em dois componentes predominantes, um tipo oeste africano e um tipo banto, que confirmam o que é conhecido dos registros sobre o tráfico de escravos para as ilhas<sup>15</sup> (cf. HAGEMEIJER; ZAMORA, 2016: 195-196).

No regime de plantação, “o crioulo ter-se-á rapidamente difundido para as roças, (...), tornando-se a língua-alvo dos escravos recém-chegados para efeitos de comunicação”. (HAGEMEIJER, 2009: 4). Conforme se comentou na introdução deste capítulo, em decorrência, sobretudo, do fluxo migratório para outras regiões, iniciou-se a ramificação no espaço e no tempo do PCGG e

as variedades desta língua de contato foram levadas para as ilhas adjacentes do Príncipe, onde se tornou conhecida como Lung’iê ‘língua da ilha’ (Principense), e Ano Bom, onde evoluiu para o Fa d’ambô ‘falar de Ano Bom’, e também originou o Angolar (Ngola, Lunga Ngola), um crioulo quilombola usado em várias partes de São Tomé.<sup>16</sup> (HAGEMEIJER, no prelo)

<sup>14</sup> Tradução de “The shift towards a plantation Society corresponds to the emergence and subsequent dominance of Banto areas, first the Congo and then Angola, in the slave trade to Sao Tome”.

<sup>15</sup> Tradução de “Genetic studies that were carried out on São Tomé and Príncipe show the admixture of the populations that became the inhabitants of the islands over the centuries. While the European genotype represents only some 10 percent of the São Tomé genotype, the African genotype makes up for the other 90 percent and can be further refined into two predominant components, a West-African and a Banto type, which confirms what is known from the records of the slave trade to the islands”.

<sup>16</sup> Tradução de “varieties of this contact language were taken to the neighboring islands of Príncipe, where it became known as Lung’Ie ‘language of the island’ (Principense), and Annobón, where it evolved into Fa d’Ambô ‘speech of Annobón, and also originated Angolar (Ngola, Lunga Ngola), a maroon creole spoken in several parts of São Tomé”.

No fim do século XVI, acentua-se a crise socioeconômica. Os cativos passaram a ser criados como animais: eram impedidos de semear para subsistência, foram separados de suas famílias e tiveram revogado o direito à liberdade dos filhos, até então decretado (cf. GARFIELD, 1992 *apud* BANDEIRA, 2017: 121). As fugas das roças para áreas distantes e a constituição de novas comunidades quilombolas, atestadas desde os últimos anos do século XV, intensificaram-se nesse momento. Colonos entraram em conflito com os escravos fugidos, no que Caldeira (2008: 50) denomina ‘guerra do mato’. Na mesma época, a ilha foi invadida por piratas e sofreu uma série de atrocidades (cf. NEGREIROS, 1895). Finalmente, a capital foi saqueada por um ataque holandês (cf. FERRAZ, 1979: 19). Sobre o momento de plena decadência de São Tomé, Cortesão (1968: 42 *apud* BANDEIRA, 2017: 122) afirma que os fazendeiros com quem os combatentes angolares se preocupavam desde

1575, e os comerciantes e os construtores de barcos, continuamente perseguidos no mar, tinham a partir do final do século XVI abandonado a ilha em amplos números, a maioria deles ia para o Brasil. A indústria açucareira declinou notavelmente; dos muitos engenhos não mais que ruínas permaneciam; e parte da cidade antiga e densamente povoada agora estava deserta e destruída.

O êxodo europeu contribuiu para o desenvolvimento e para a difusão das línguas crioulas: com uma referência do Português cada vez mais limitada, a protolíngua parece buscar influências lexicais e fonológicas nas línguas de substrato africano (cf. FERRAZ, 1979: 19-20). Conforme se mencionou na seção 1, apesar de partilharem uma origem comum, hoje, a inteligibilidade entre tais crioulos é limitada, sobretudo “pela difusão do PCGG no tempo e no espaço e pelo subsequente desenvolvimento de cada variedade”<sup>17</sup>, através de novos contatos secundários (HAGEMEIJER, 2011: 112). Três séculos depois da descoberta, o país se torna, de novo, palco de uma série de mudanças socioeconômicas, cujos reflexos sociolinguísticos perduram até os dias atuais e serão comentados a seguir.

## 2.2 A segunda colonização e os ciclos do café e do cacau

No fim do século XVIII e primórdios do XIX, inauguram-se os respectivos ciclos do café e do cacau. Trata-se de um período de crise trabalhista, acentuado pela abolição da escravatura (1869) e pela formalização da condição jurídica dos libertos (1875). Uma vez livres, eles se negavam a trabalhar nas empresas agríco-

---

<sup>17</sup> Tradução de “by the diffusion of the proto-GGC in time and space and the subsequent development of each variety”.

las e a população local não supria a demanda de mão-de-obra. Na tentativa de garantir a força de trabalho, os europeus criaram a ‘Curadoria Geral dos Indígenas’, “solução, subjacente a uma revolução social e linguística inigualável nas ilhas através do recrutamento massivo de trabalhadores contratados (serviçais ou contratados) principalmente de outras colônias portuguesas em África”<sup>18</sup> (HAGEMEIJER, no prelo).

O regime de contrato, instaurado em 1875, representa um novo marco sociolinguístico, sobretudo pelo fluxo migratório de Cabo Verde, Angola, Moçambique, Benim e Gabão para São Tomé. Dados estatísticos de Nascimento (2000, *apud* GONÇALVES, 2016) atestam, a partir de então, um aumento demográfico notável, decorrente, em especial, da chegada de serviçais e de mais portugueses. Para além do PST e dos crioulos autóctones, suas línguas maternas (dentre as quais sobressaem o crioulo de Cabo Verde e variedades bantas como o Quimbundo e o Umbundo) ampliaram o multilinguismo. Segundo Hagemeyer (no prelo), “tal como os escravos africanos que chegaram em São Tomé e Príncipe no período inicial, no começo os trabalhadores contratados não falavam Português”<sup>19</sup>, mas, diferentemente do século XVI, nas senzalas das plantações, se mantinham segregados dos nativos em um ambiente exclusivo sociocultural e linguisticamente (cf. SEIBERT, 2014: 61). Trabalhando seis dias por semana, eles

tinham pelo menos alguma exposição ao Português, a língua dominante dos plantadores, cujo número aumentou desde o boom do café e do cacau. Os portugueses constituíam o principal grupo de proprietários, pois o governo santomense negou aos forros donos de terras o acesso aos trabalhadores contratados. (...) A população livre, em particular o grupo maior, os Forros, que viviam nas matas e vilas, fora dos novos sistemas coloniais, se mantem (deliberadamente) segregada dos serviçais, que eles consideram os ‘novos escravos’ e, portanto, de status inferior<sup>20</sup> (HAGEMEIJER, no prelo).

<sup>18</sup> Tradução de “a solution which underlies an unparalleled social and linguistic revolution on the islands through the massive recruitment of contract workers (*serviçais* or *contratados*) mainly from other Portuguese colonies in Africa”.

<sup>19</sup> Tradução de “Just like the African slaves that arrived in São Tomé e Príncipe in the earlier period, in the beginning the contract laborers didn’t speak Portuguese”.

<sup>20</sup> Tradução de “had at least some exposure to Portuguese, the dominant language of the planters, whose number had also gone up since the coffee and cocoa boom (...) the Portuguese constituted the main group of land owners, since the Santomean government denied Forro landowners access to contract workers (,,,) the free population, in particular the largest group, the Forros, who lived in the forests and villages, i.e. outside the new colonial system, kept (deliberately) segregated from the contract laborers, which they considered the ‘new slaves’ and therefore of lower status”.

O Português tornou-se a referência naquele ambiente e “a L2 que os contratados começariam tipicamente a adotar”<sup>21</sup>. Surgiu, portanto, a variedade dos Tongas, fruto do contato entre as línguas maternas dos recém-chegados e aquela empregada nas roças. De acordo com Gonçalves (2016: 25), é provável que suas múltiplas formas “tenham emergido em função das idiosincrasias das diferentes roças em que eram faladas”. Para Lopes; Baxter (2011a: 2), as variedades foram adquiridas “no trabalho em São Tomé, principalmente a partir de modelos de português L2 falados por trabalhadores colegas e, indiretamente, do português L1 dos administradores”. Devido ao acesso limitado da maioria dos habitantes à língua e à “existência de poucas escolas primárias para a população crioula”<sup>22</sup>, adotam-se medidas como o decreto de 1878, que exige dos plantadores a disponibilização “de educação para seus funcionários e para os filhos deles”, (HAGEMEIJER, no prelo). Representam, entretanto, “apenas uma gota no oceano e não refletem uma política linguística consistente pró-Português”<sup>23</sup> (HAGEMEIJER, no prelo).

Conforme será reiterado na seção 3.2, a transição da Língua Portuguesa de L2 para L1 é um fenômeno subsequente à emancipação, quando o idioma é escolhido como língua oficial exclusiva, por já possuir uma norma padrão e ajudar a garantir a unidade nacional em meio à diversidade etnolinguística. Rapidamente, “deixou de ser a língua da elite e dos domínios mais altos ou formais, se tornando a língua de todos os contextos comunicativos e a L1 da maioria dos habitantes do país”<sup>24</sup>. Dentre os fatores influentes na mudança estão: i) o direito universal à educação (em Português); ii) a mobilidade social promovida pelo fim da era colonial; iii) o maior contato com meios de comunicação (rádio e televisão); e iv) a ausência de políticas linguísticas pró-crioulas (cf. GONÇALVES; HAGEMEIJER, 2015: 91). Por conta de todas as consequências da independência, não resta “muito do português dos tongas, mas o seu legado linguístico contribuiu para o que é hoje a variedade do português de São Tomé e Príncipe.” (GONÇALVES; HAGEMEIJER, 2015: 90).

As seções futuras concentram-se no caráter plurilíngue da ilha, abordando tanto a evolução do PCGG nos quatro sistemas autóctones, quanto a disseminação do PST como língua veicular.

---

<sup>21</sup> Tradução de “the L2 that the contract laborers would typically start to adopt”.

<sup>22</sup> Tradução de “the existence of few primary schools for the creole population”.

<sup>23</sup> Tradução de “just a drop in the ocean and do not reflect a consistent pro-Portuguese language policy”.

<sup>24</sup> Tradução de “ceased to be the language of the elite and of the high or formal domains, becoming the language of all communicative contexts and the L1 of the majority of the country’s inhabitants”.

### 3. O MULTILINGUISMO EM SÃO TOMÉ E PRÍNCIPE

#### 3.1 Os Crioulos do Golfo da Guiné

##### a) O Forro ou Santome

Como se comentou na seção 2.1, a gênese do Santome, também chamado de Forro, remete à constituição de uma nova sociedade. Trata-se de um estrato social originalmente composto pelos primeiros escravos libertos após os decretos de Dom Manuel, em 1515 e 1517. Adquirindo direitos e poderes próprios, tal comunidade teria interferido, não só na origem, mas também na consolidação da nova variedade que, aos poucos, se expandiu pela sociedade colonial emergente (cf. HAGEMEIJER, 2009: 5). Mais que o produto do contato entre o Português (superstrato) e línguas Edo (substrato), o Santome ascendeu como uma língua de resistência e identidade social, durante a fase de habitação. Nas palavras de Lorenzino (1996: 435 *apud* BANDEIRA, 2017: 128), seu prestígio justifica-se por representar os “mestiços que atingiram um influente *status* socioeconômico quando se converteram em proprietários de terras e escravos”. Não se pode, porém, descartar a existência de

um certo nível de bilinguismo entre os forros integrantes das classes sociais mais altas e entre parte da população escrava que tinha frequente contato direto com os Portugueses, tais como os escravos de casa (opondo-se aos escravos de plantação) e as mulheres cativas e seus descendentes como resultado de concubinatos.<sup>25</sup> (HAGEMEIJER, no prelo)

Faltam dados exatos da distribuição das línguas autóctones pela população na era colonial, pois os recenseamentos só começam a considerá-las, após a independência do país, em 12 de julho de 1975. Estudos defendem, todavia, o acesso limitado ao Português e o predomínio dos crioulos. Com base em alguns testemunhos da época, Gonçalves; Hagemeijer (2015: 90) concluem:

Pelo menos até finais do século XIX, as línguas crioulas dominam a paisagem linguística das ilhas. A utilização do português parece estar circunscrita aos contextos que requerem uma relação mais direta com o regime colonial, havendo, em princípio, poucos falantes nativos desta língua.

---

<sup>25</sup> Tradução de “a certain degree of bilingualism among the Forros belonging to the higher social classes and among parts of the slave population which had frequent direct contact with the Portuguese, such as the house slaves (as opposed to the plantation slaves) and female slaves and their descendants as the result of concubinages”.

Em outras palavras, o contato com a língua do colonizador restringia-se aos santomenses mais escolarizados, integrantes ou aspirantes à elite urbana. Por outro lado, para a população analfabeta, residente nas áreas periurbanas e rurais, os crioulos representavam as “línguas de comunicação quotidiana” (cf. MATA, 2010: 17). Nesse cenário, a variedade dos alforriados, adquiriu um *status* social privilegiado e constituiu, desde o século XVI, a língua-alvo dos escravos recém-chegados.

A comercialização do açúcar aumenta a demanda por mão-de-obra e, para além da zona de Benim, a ilha recebeu escravos do Congo e de Angola, falantes de línguas banto, como o Quicongo e o Quimbundo. Pautado em indícios de que os quatro crioulos se estabilizaram até 1520, quando o Edo era mais representativo na população, Hagemeyer (2009: 16-17) adverte: “nestas condições, o quimbundo não deve ser considerado uma língua de substrato, um papel reservado ao Edo, mas sim uma língua de adstrato, sem efeito fundador”.

A primeira alusão à “Lengua de San Thomé” data do século XVII: em 1627, o Padre Alonso de Sandoval, recém-chegado na ilha, menciona sua existência. Em 1766, Gaspar Pinheiro da Câmara (*apud* HOLM, 1988) também a cita. No entanto, somente no século XIX, divulgam-se estudos sobre a temática: Schuchardt o faz em 1882 e em 1895, Negreiros publica, na *Historia Ethnographica da Ilha de São Thomé*, um capítulo a respeito (cf. HAGEMEIJER, 2000).

Segundo Neves; Ceita (2004 *apud* HAGEMEIJER, no prelo), no fim do século XVIII, os forros representavam 95% da população de escravos libertos (30% da população total de São Tomé) e até o fim do século XIX, sua língua nativa se manteve como a mais falada na região. Com o regime de contrato e a independência, o domínio diminuiu rapidamente e cada vez menos é adquirida como L1. Em contrapartida, o Português, variedade mais prestigiada, se expandiu (cf. seções 2.2 e 3.2), tornando-se o idioma veicular da maioria da população urbana<sup>26</sup>. Entre os sistemas autóctones, contudo, preserva-se o *status* privilegiado do Santomé. No interior do país, é muito frequente, associado à tradições e canções do grupo Forro. Dessa forma, atualmente, seu uso sucede o português: 36% (62.707) dos 173.015 habitantes do país, declaram falá-lo (RGPH, 2012).

#### b) O Angolar

No que remete à formação da comunidade angolar, três hipóteses competem. Segundo Seibert (2004), a primeira delas foi publicada no início do século

---

<sup>26</sup> De acordo com o Factbook, (2009 *apud* BANDEIRA, 2017: 128), estima-se que 61% da população santomense residam nas áreas urbanas, ao passo que 39% se concentrem nas áreas rurais do país.

XVIII e sugere serem sobreviventes de um naufrágio que acometeu um navio de escravos trazidos de Angola, no século XVI. A segunda defende que São Tomé não era desabitada quando os portugueses chegaram, isto é, a comunidade angolana original corresponderia aos habitantes autóctones da ilha. A terceira possibilidade, enfim, aponta a fuga de escravos para áreas até então desconhecidas. Esses fugitivos teriam formado o primeiro grupo, ampliado nos séculos XVI e XVII, após receber outros advindos das roças e das cidades com propósitos semelhantes.

Embora a hipótese de fuga dos cimarrones (cf. SEIBERT, 2004) ou quilombolas (cf. BANDEIRA, 2017) seja a mais recente, muitas evidências linguísticas e históricas parecem corroborá-la. Para Ferraz (1979), trata-se da possibilidade mais plausível, pois estudos comparativos comprovam a proximidade entre Angolar, Forro e Lung'ie. Há evidências de que os angolares falavam uma língua de base portuguesa e não banto, apesar de não se descartar uma possível aculturação histórico-linguística. Partindo de tal pressuposto, o autor questiona: como se justificaria a influência se estivessem isolados do restante da população, sendo sobreviventes de um naufrágio ou os primeiros habitantes do país? A necessidade de estudos mais específicos para infirmar ou confirmar as outras proposições tem promovido, portanto, a tese de que seriam escravos fugidos depois da formação do PCGG, cuja língua apresenta traços dessa variedade comum com especificidades derivadas do isolamento posterior.

Corroborando a perspectiva de Ferraz (1979), diversos autores ressaltam a existência de documentos históricos descrevendo a recorrência, à época, de evasões de escravos para o interior da ilha. Em 1499, Álvaro de Caminha comenta essas saídas para áreas campestres mais afastadas da capital (cf. ALBUQUERQUE, 1989 *apud* SEIBERT, 2004: 55). Migravam para espaços montanhosos e de difícil acesso, cansados dos maus tratos sofridos, das péssimas condições de trabalho oferecidas, da escassez de alimentos e do cerceamento à liberdade. Segundo Dias; Diniz (1988 *apud* BANDEIRA, 2017: 134), tal ambiente favoreceu a construção de quilombos e, durante os séculos XVII e XVIII, a inacessibilidade e os declínios demográfico e econômico garantiram relativo isolamento até a última parte do século seguinte.

Complementando tal ponto de vista, Bandeira (2017: 135) adverte:

não se pode afirmar que os quilombolas viviam em isolamento completo. Os assaltos à capital somados ao sequestro de mulheres das fazendas evidenciam que falantes de angolar tiveram contato com falantes de santome, o que viabilizou a troca de elementos linguísticos entre ambas as línguas.

Segundo a autora, há relatos de ataques de escravos à cidade (cf. HENRIQUES, 2000: 116) e de fazendas abandonadas pelos colonos após ameaças de

foragidos (cf. SANTOS, 1996: 81). Nas zonas oeste e sul de São Tomé, mais próximas dos quilombolas, mesmo as expedições militares não garantiam a reconquista dos colonos (cf. CALDEIRA, 2004: 113). Em 1693, inclusive, o capitão do mato, Mateus Pires, comandou a chamada pacificação, uma última tentativa infrutífera de controlar os fugitivos sequestradores de mulheres na capital e nas plantações. De acordo com Almeida (1962: 10), desde então, não se encontram mais evidências de invasões aos espaços dos foragidos, sob pena de morte.

Hoje, os angolares distribuem-se em comunidades situadas no litoral, tanto no distrito de Caué, entre Ribeira Afonso e Porto Alegre, quanto no distrito de Lembá, entre Neves e Bindá. Mais próximos de São Tomé, há, ainda, pequenos grupos de falantes na Praia Melão, Pantufo e São João da Vargem (cf. CEITA, 1991 *apud* BANDEIRA, 2017: 135). Resultados do último censo estimam que 6% (11.377) da população total do país (173.015) falem esse crioulo (RGPH, 2012). Com base nesses números, Hagemeyer (no prelo), afirma

A comunidade Angolar, que é tradicionalmente vista como um grupo mais homogêneo, também está sofrendo hoje um impacto massivo do Português, mas foi exposta previamente ao Forro, devido a sua difusão ao longo das áreas costeiras (...). Mesmo em seus redutos tradicionais, no distrito de Cauê, quase 100% da população afirma falar o Português, enquanto o Angolar é falado apenas por aproximadamente 50% nesta região. (...) é na faixa etária de 20 a 24 anos que mais de 50% ainda falam o Angolar (RGPH, 2012). Note-se, ainda, que os falantes do Angolar estão espalhados proporcionalmente por todos os distritos, exceto Lobata e Príncipe, onde seus números são menores, o que também contribui para o comprometimento desta língua.<sup>27</sup>

### c) *O Lung'ie ou língua do Príncipe*

Pouco depois de aportarem em São Tomé, também na década de 1470, os portugueses chegam ao Príncipe (cf. seção 1). O nome original da Ilha, Santo Antão, mais tarde foi alterado para o atual em homenagem ao Príncipe Perfeito,

<sup>27</sup> Tradução de “the Angolar community, which is traditionally seen as a more homogenous group, is nowadays also undergoing the massive impact of Portuguese, but had previously already been exposed to Forro due to their spread along (...) coastal areas. Even within their traditional stronghold, the district of Caué, almost 100% of the population claims to speak Portuguese, whereas Angolar is only spoken by roughly 50% in this region. (...) it is from the age group of 20-24 on that over 50% still speaks Angolar (RGPH, 2014). Note further that the Angolar speakers are spread out quite proportionately over all the districts, except for Lobata and Príncipe, where their numbers are lower, which also contributes to the endangerment of this language.”.

Dom João. Em 1500, trinta anos após a chegada dos colonos, o território é doado por decreto real e se inicia o processo de colonização (cf. HENRIQUES, 2000; MAURER, 2009 *apud* BANDEIRA, 2017: 129), quando o PCGG, já constituído, teria sido levado para lá. Distante e isolada do restante do arquipélago, a língua-mãe estabelece contatos com outras variedades trazidas por novos escravos e, paulatinamente, se desenvolve a língua do Príncipe ou Lung'ie.

O isolamento parece implicar reflexos importantes: se comparado aos demais crioulos do Golfo da Guiné, no Lung'ie, haveria “mais léxico de origem edo e (...) aspectos fonológicos edóides, exclusivos da área onde essas línguas eram faladas” (BANDEIRA, 2017: 130). Em outras palavras, nesse sistema, a influência linguística do Delta do Níger se destaca, por i) representar a ramificação mais antiga do proto-crioulo do Golfo da Guiné; ii) ter sido menos afetada pelo ciclo de plantação e pela influência da força de trabalho banto dele decorrente; e iii) ter garantido, entre 1514 e 1518, o monopólio do comércio português com o Reino de Benim, preservando a ligação com o Delta do Níger (cf. HAGEMEIJER, 2011 a: 114).

Por volta de 1900, a população começa a declinar: a ilha do Príncipe sofre uma epidemia de tripanossomíase (doença do sono) e apenas 300 pessoas sobrevivem. Em virtude do impacto demográfico significativo e da necessidade de mão-de-obra, importaram-se trabalhadores contratados de diferentes áreas, como Angola, Moçambique, São Tomé e Cabo Verde (cf. MAURER, 2009).

Tendo em vista fatores como a quantidade de crianças que adquirem a língua-alvo como materna; a atitude da população perante a língua; e o impacto de outras línguas em potencial ou real competição (cf. CRYSTAL, 2000 *apud* BANDEIRA, 2017: 131), pode-se dizer que, atualmente, a língua do Príncipe está em risco de extinção. Com base nesses parâmetros, Crystal (2000) sugere três níveis de vitalidade linguística: seguro, ameaçado e extinto.<sup>28</sup> Sob os critérios do autor, o Lung'ie estaria ameaçado, por conta da escassez de falantes (crianças) que o adotam como L1 e do “nível de impacto de outras (...), como o português, língua oficial e mais utilizada em todas as esferas sociais, e o kabuverdianu, haja vista que há muitos falantes descendentes dos trabalhadores contratados que chegaram na ilha no final do século XIX e começo do XX” (BANDEIRA, 2017: 131-132).

Para Maurer (2009: 3-4) o processo de extinção decorre: i) da “epidemia da doença do sono, cuja consequência foi a de que os falantes nativos do lung'ie foram rapidamente superados em números pelos trabalhadores contratados”; ii) da

---

<sup>28</sup> A depender das escalas adotadas para determinar a vitalidade de uma língua, níveis diferentes podem ser encontrados. Nas escalas da UNESCO e dos linguistas Joshua Fishman e Paul Lewis, apresentam-se, 6, 8 e 13 níveis diferentes de extinção, respectivamente.

falta de transmissão da língua “para os mais jovens por três ou quatro gerações”; e iii) de nenhuma estratégia ter sido adotada, anteriormente, na tentativa de promover o acesso dos imigrantes, sobretudo de Cabo Verde, a ela. Esses e outros fatores conduziram a uma maior difusão do kaboverdiano, depois de variedades regionais do Português e do Forro.

O último censo aponta cerca de 1.753 sobreviventes da comunidade, superando claramente os números apresentados em outras fontes: enquanto Maurer (2009: 3) afirma haver apenas 20 falantes ativos, Agostinho (2015) defende a existência de cerca de 200 utentes, sendo a maioria idosos bilíngues cuja fluência é variável. Segundo Hagemeyer (no prelo), trata-se de “estimativas (...) baseadas exclusivamente na situação linguística do Príncipe. Para esta ilha, o censo de 2012 considera 525 falantes, ao passo que os 1.228 restantes foram contados em São Tomé (a maioria em Água Grande, o distrito da Capital)”<sup>29</sup>. Ademais, “a aparente superestimação do número de falantes pode estar parcialmente relacionada a atitude linguística e não apenas ao uso da linguagem.”<sup>30</sup>

#### d) *O Fa d’ambô ou falar de Ano Bom*

Como foi esclarecido na seção 2, o Fa d’ambô (ou falar de Ano Bom) é uma das línguas filhas do protocrioulo do Golfo da Guiné, fruto da migração para Ano Bom, uma das províncias da Guiné Equatorial, situada ao sudoeste de São Tomé. A ausência de evidências sobre a data exata da descoberta impede o consenso entre os teóricos, mas, em consonância com Caldeira (2010 *apud* HAGEMEIJER; ZAMORA 2016: 195), parece plausível situar o evento entre 1483 e 1501, no primeiro dia do ano, devido à tradição portuguesa de batizar áreas encontradas de acordo com o calendário. Em 1503, Jorge de Melo é nomeado capitão donatário e realiza um povoamento pioneiro também no começo do século XVI. Até pelo menos 1518, a ilha atuou como entreposto de escravos, e documentos comprovam a existência de tráfico marítimo ali ainda nos anos 1530, quando foi abandonada (cf. HAGEMEIJER; ZAMORA, 2016: 196). A colonização permanente deve ter ocorrido, enfim, por volta de 1543 e 1565, respectivos momentos dos registros de que i) a região estava deserta; e ii) fora habitada por um homem branco e alguns escravos plantadores de algodão (cf. CALDEIRA, 2010). O atraso seria decorrente de fatores como a “pequena extensão territorial (...), o relevo montanhoso que deixava poucas áreas férteis cultiváveis, (...) a falta de

<sup>29</sup> Tradução de “Estimates (...) based exclusively on the language situation on Príncipe. For this island, the 2012 census counted 525 speakers, whereas the remaining 1,228 were counted on São Tomé (most of which in Água Grande, the district of the capital).

<sup>30</sup> Tradução de “The apparent overestimation of the number of (active) Lung’le speakers might be partly related to language attitude and not just language use”.

ancoradouros seguros, sua posição longe das rotas comerciais portuguesas e (...) a forte concorrência com a ilha de São Tomé” (BANDEIRA, 2017: 136).

Face às dificuldades de se adotar um regime de produção agroindustrial, Portugal aderiu à partilha. Nela, o capitão donatário delegava funções ao clero local ou aos comerciantes de escravos (cf. CALDEIRA, 2010 *apud* HAGEMEIJER; ZAMORA, 2016: 196), dentre as quais se destacava uma importante fonte de renda: “a entrega de uma quantidade de algodão, às vezes, já tecido em faixas” (BANDEIRA, 2017: 136). Mesmo detendo o controle, a presença europeia na região era mínima, restringindo-se, por um longo período, ao representante do capitão-donatário, à guarda pessoal e ao feitor. A maioria da população era, assim, composta por escravos inicialmente trazidos “de São Tomé e (...) do Príncipe, ou (...) diretamente do continente africano, em particular do Congo e da Angola”<sup>31</sup> (HAGEMEIJER; ZAMORA, 2016: 196). Nesse cenário, “o crioulo levado a Ano Bom, durante o pico da economia de plantação, pode ter sido uma forma do Proto-crioulo do Golfo da Guiné (= Santomense) que já tinha sofrido influência do estrato Banto Ocidental”<sup>32</sup> (HAGEMEIJER, 2011 a: 114).

Reforçam a hipótese de uma origem santomense evidências da migração de falantes de uma língua dominante ou mesmo nativa para Ano Bom no fim do século XVI e início do XVII (cf. HAGEMEIJER; ZAMORA, 2016: 199-200). Em 3 de novembro de 1770, o capitão donatário de São Tomé e Príncipe, após acompanhar Antônio Luís Monteiro e Gregório Martins das Neves em uma viagem a Ano Bom, escreve uma carta em que relata a semelhança entre sua língua local e aquela falada pelos habitantes da ilha. Tal conexão é reafirmada em Matos (1842: 107 *apud* HAGEMEIJER; ZAMORA, 2016: 199), ao expor que “[...] o dialecto (...) de Anno Bom é o mesmo que o de S. Thomé, mas com uma pronúncia gutural semelhante a dos Árabes”<sup>33</sup>.

<sup>31</sup> Tradução de “from São Tomé, and (...) also from Príncipe, or (...) directly from the African mainland, in particular the Congo and Angola”.

<sup>32</sup> Tradução de “the creole taken to Annobon, during the peak of the plantation economy, must have been a form of the proto-GGC (= ST) that had already undergone influence from the western Banto layer.”

<sup>33</sup> Embora a analogia entre as línguas autóctones do Golfo da Guiné seja antiga, os primeiros estudos sobre crioulos surgem no fim do século XIX e se intensificam ao longo do XX. Em um trabalho pioneiro sobre a relação entre o Fa d’ambô e o Santome, Schuchardt (1888) comprova a gênese daquele nesse. Em meados do século XX, Barrena (1957) apresenta informações estruturais sobre a língua, em uma gramática póstuma. Desde então, vários estudos têm ampliado o conhecimento sobre o falar em foco (cf. VALKHOFF, 1966; GRANDA, 1985, 1986; FERAZ, 1979; POST, 1998; ZAMORA, 2009, 2010; SILVEIRA, 2013; HAGEMEIJER; ZAMORA, 2016).

Distante dos europeus, “o contato [da população anobonense] com o mundo exterior limitava-se ao fluxo de navios em busca de água doce e de mantimentos”<sup>34</sup> (HAGEMEIJER; ZAMORA, 2016: 196). Na passagem dos séculos XVII e XVIII, se extingue, pois o último representante português deixa a ilha, expulso ou morto (cf. CALDEIRA, 2008: 3). Consequentemente, por quase duzentos anos, todos os imigrantes foram impedidos de lá entrar, garantindo a liberdade e o isolamento da população de Ano Bom. Sem o “input português” e o “afluxo de novos escravos”, a língua evoluiu livremente (BANDEIRA, 2017: 136), classificando-o como um crioulo fundamental para a compreensão da formação e evolução das demais línguas autóctones de base portuguesa, não só no Golfo da Guiné, mas também em outras áreas do Atlântico.

Em 1778, pelo Tratado de El Pardo, Portugal cede o Rio Muni e as ilhas de Ano Bom e Fernando Po à Espanha. Nesse capítulo de sua história,

a conexão com São Tomé se torna uma característica do passado e seu destino seria agora decidido por Santa Isabel (Malabo, a capital de Fernando Po (Bioko). No entanto, a nova conexão foi apenas ativada no fim do século XIX, pela chegada da missão Claretiana na ilha<sup>35</sup> (HAGEMEIJER; ZAMORA, 2016: 193).

Segundo Hagemeyer; Zamora (2016: 202), atualmente, há cerca de 5.600 falantes de Fa d’ambô: 5.000 distribuídos por Ano Bom, Bioko e Malabo, na Guiné Equatorial; e 600 pela diáspora, a maioria na Espanha. Em Ano Bom, estimativas apontam uma comunidade linguística cujo número varia entre 2.000 (Ethnologue) e 3.000 (Zamora, 2009) utentes; já em Bioko, grande parte concentra-se nas proximidades da capital Malabo, estendendo-se a algumas vilas da costa e à cidade de Bata. Embora não existam registros oficiais,

pode-se seguramente assumir que o número de falantes de Fa d’ambô como L2 é muito pequeno. A língua não é usada como uma língua franca em outras comunidades na Guiné Equatorial e existe uma tendência para os anobonenses se casarem no âmbito de sua própria comunidade de fala. Além do Espanhol como língua oficial, utentes do Fa d’ambô, especialmente aqueles que vivem em Malabo, são usualmente falantes do Pichi também. As duas

---

<sup>34</sup> Tradução de “contact with the outside world was limited to passing vessels in search of fresh water and supplies”.

<sup>35</sup> Tradução de “Annobón’s connection with São Tomé became a feature of the past and its destiny would now be decided from Santa Isabel (Malabo), the capital of Fernando Po (Bioko). However, (...) new connection was only activated at the end of the nineteenth century, through the arrival of the Claretian mission on the island”.

linguagens têm deixado marcas no crioulo, que são essencialmente circunscritas ao domínio lexical<sup>36</sup> (HAGEMEIJER; ZAMORA, 2016: 202).

Trata-se de uma das menores comunidades de fala da região da Guiné, cujos integrantes se distribuem por diferentes espaços. A língua é, porém, “transmitida entre gerações e goza de prestígio social entre seus falantes”<sup>37</sup> (HAGEMEIJER; ZAMORA, 2016: 202). A propagação e o prestígio parecem resultar de uma história de isolamento, legitimada na maior resistência às influências externas e “contribuem para um forte espírito de solidariedade e auto-identidade”<sup>38</sup>. Além disso, o difícil relacionamento entre a minoria anobonense e o grupo étnico dominante na região, o Fang, também “poderia ser considerado um fator que promove a identidade do grupo e a manutenção da linguagem”<sup>39</sup> (HAGEMEIJER; ZAMORA, 2016: 203).

De acordo com Granda (1985 *apud* HAGEMEIJER; ZAMORA, 2016: 203), até a independência da Guiné Equatorial, há pouca exposição do crioulo à linguagem oficial, pois o uso do espanhol limitava-se aos missionários e a alguns professores e oficiais. No início dos anos 1980, enquanto o conhecimento do espanhol começa a se difundir pelos homens adultos, mulheres e crianças preservam-se como monolíngues em Fa d’ambô. Ademais, como os homens mais jovens e de meia idade passam grande parte de seu tempo trabalhando em Bioko, usam o Pichi com seus pares. Em consonância com Post (1998), Hagemeijer; Zamora (2016: 203) afirmam que o Fa d’ambô é a língua nativa na ilha de Ano Bom, mas a maioria da população valoriza e fala o idioma oficial. Existe, no Bioko, um grupo de imigrantes (homens e mulheres jovens e de meia idade) aportados para estudar ou trabalhar, que só retornam à ilha em uma idade mais avançada, caracterizando-a como “um espaço para crianças e idosos. Além de sua língua nativa, o espanhol é a língua exclusiva usada na educação primária e secundária”<sup>40</sup> (HAGEMEIJER; ZAMORA, 2016: 203). Assim, devido ao “maior isolamento e à

---

<sup>36</sup> Tradução de “it can safely be assumed that the number of L2 Fa d’Ambô speakers is very low. The language is not used as a lingua franca by other speech communities in Equatorial Guinea and there is a tendency for Annobonese to marry within their own speech community. In addition to the official language Spanish, Fa d’Ambô speakers, especially those who live in Malabo, are usually Pichi speakers as well. Both these languages have left imprints on the creole, which are mostly restricted to the lexical domain”.

<sup>37</sup> Tradução de “transmitted across generations and enjoys social prestige among its speakers”.

<sup>38</sup> Tradução de “contributed to a strong spirit of solidarity and self-identity”.

<sup>39</sup> Tradução de “should be considered a factor that promotes group identity and language maintenance”.

<sup>40</sup> Tradução de “a place for children and the elderly. In addition to their native language, Spanish is the exclusive language used in primary and secondary education”.

idade da comunidade em Ano Bom, espera-se encontrar uma variedade mais conservadora ali”<sup>41</sup> (HAGEMEIJER; ZAMORA, 2016: 204).

Até o início dos anos 1990, o território permanece isolado, isento de influências midiáticas alheias, como rádio, televisão ou livros. Em contrapartida,

na capital Malabo, (...) falantes do Fa d’ambô se expõem a outras línguas, em particular ao espanhol e à língua franca Pichi, falada por alguns. Apesar da coexistência dessas línguas, (...) não houve variação significativa entre os falares de Ano Bom e Malabo devido à contínua migração entre os dois espaços<sup>42</sup> (HAGEMEIJER; ZAMORA, 2016: 203).

Vinte anos depois, esse cenário evolui: em 2010, Ano Bom inaugura um porto e um aeroporto e disponibiliza o acesso aos meios de comunicação midiáticos. Com base nessas mudanças, Hagemeijer; Zamora (2016: 203) salientam: “enquanto o Fa d’ambô não está incluído na lista de linguagens ameaçadas da UNESCO, os sinais de que existem pontos fracos na transmissão geracional (...) não podem ser desconsiderados.”<sup>43</sup> Atualmente, a língua está perdendo sua fluidez: existem crianças anobonenses que não falam o crioulo (cf. ZAMORA, 2010 *apud* HAGEMEIJER; ZAMORA, 2016: 204) e jovens em Malabo frequentemente usam o Pichi para se comunicarem (cf. HAGEMEIJER; ZAMORA, 2016: 204). Portanto,

a mudança para o Pichi, que é mais proeminente entre as gerações mais jovens (...) poderia ser identificada como a principal ameaça ao Fa d’ambô. Comparada ao espanhol, a língua franca formal, o Pichi prospera como língua franca informal, a língua das estradas, do comércio e mercados, afetando a estável diglossia que caracteriza a relação entre Fa d’ambô e o espanhol. O Pichi é o competidor direto do Fa d’ambô na esfera informal do Bioko<sup>44</sup> (HAGEMEIJER; ZAMORA, 2016: 204).

<sup>41</sup> Tradução de “the more isolated and aged community on Annobón, the more conservative variety is expected to be found there”.

<sup>42</sup> Tradução de “in the capital Malabo, (...) Fa d’Ambô speakers were exposed to other languages, in particular to Spanish and to the lingua franca Pichi, spoken by many. Despite the coexistence of these languages, (...) there was no significant variation between Fa d’Ambô spoken on Annobón and in Malabo due to ongoing migration between both spaces”.

<sup>43</sup> Tradução de “While Fa d’Ambô is not included on the UNESCO list of endangered languages, the signs that there are weak spots in the generational transmission (...) should not be dismissed.”

<sup>44</sup> Tradução de “This shift to Pichi, which is more prominent among the younger generations (...) should be identified as the main threat to Fa d’Ambô. Compared to Spanish, the formal lingua franca, Pichi thrives as the informal lingua franca, the language of the streets, commerce, and markets, affecting the stable diglossia that characterizes the relation between

Convive com tais línguas, na ilha de Ano Bom, uma variedade de discurso litúrgico baseada no Português (arcaico) (...) “que compreende elementos lexicais e gramaticais do português, do Fa d’ambô e provavelmente também umas poucas expressões do Latim”<sup>45</sup> (HAGEMEIJER; ZAMORA, 2016: 200). A sobrevivência dessa “forma crioulaizada do português” parece se justificar em uma forte tradição católica introduzida em São Tomé e estimulada pelo clero, cuja presença remete “ao desenvolvimento de um culto autossuficiente, transmitido pelos sacerdotes locais que atuam em paralelo às igrejas para satisfazer as necessidades espirituais da população”<sup>46</sup> (HAGEMEIJER; ZAMORA, 2016: 201).

Em suma, apesar de se reconhecer que a relação entre o Fa d’ambô e a identidade anobonense contribui para a preservação do crioulo, fatores como ser “uma pequena comunidade linguística que se tornou mais dispersa, devido à (e) migração; a crescente exposição à língua franca Pichi; e a falta de suporte governamental”<sup>47</sup> parecem colocá-lo em um crescente risco de extinção. Nas palavras de Hagemeijer; Zamora (2016: 206),

as permanentes emigração e exposição ao Pichi e ao Espanhol nos domínios baixo e alto, (...) reduzirão a participação de nativos, falantes competentes de Fa d’ambô e podem ultimamente levar a sua remodelação, por exemplo através da relexificação parcial, harmonização estrutural ou, na pior das hipóteses, à mudança linguística.<sup>48</sup>

### 3.2 A língua companheira do império em São Tomé e Príncipe

A expansão do Português para além de Portugal se inicia no ano de 1415, quando colonizadores se lançam ao mar e aportam em diferentes paragens com

---

Fa d’ambô (L) and Spanish (H). Pichi is Fa d’Ambô’s direct competitor in the informal sphere on Bioko”.

<sup>45</sup> Tradução de “comprises lexical and grammatical elements of Portuguese, Fa d’Ambô, and arguably also a few Latin expressions”.

<sup>46</sup> Tradução de “to the development of a self-sufficient cult, transmitted by local priests who functioned in parallel to the church, in order to satisfy the spiritual needs of the population”.

<sup>47</sup> Tradução de “a small speech community that has become more scattered due to (e)migration; growing exposure to lingua franca Pichi; and the lack of any type of government support for language maintenance”.

<sup>48</sup> Tradução de “the ongoing (e)migration and exposure to Pichi and Spanish in the Low and High domains (...) will reduce the share of native, competent Fa d’Ambô speakers and may ultimately lead to the reshaping of Fa d’Ambô, for instance through partial relexification, structural convergence, or in a worst-case scenario, to language shift”.

religião, cultura e, principalmente, idioma próprios. Pode-se dizer que a diacronia da língua, classificada como “companheira do império” desde Fernão de Barros e António de Nebrija, representa “a trajetória histórico-linguística de Portugal e das ex-colônias, uma história de dominação, de relações tensas, de violência exercida sobre o Outro” (AFONSO, 2009: 41).

Imbuído de uma função imperial, o Português se tornou um instrumento de submissão na relação colonizador-colonizado. No processo de expansão, era uma garantia da manutenção do império, pautada em uma ideologia de déficit, em que de um lado se tinha uma língua prestigiada, erudita, e de outro, línguas africanas inferiores, de tradição folclórica. Nesse contexto, os colonizados eram entendidos como povos não civilizados e

imperiosamente se lhes impunha romper com as milenárias estruturas mentais e sociais para se transferirem para outros planos de vida, o plano dos “assimilados”, civilizados ou cidadãos, libertos das hierarquias e disciplinas tribais, passando a viver no estilo e costumes europeus, falando, lendo e escrevendo Português. (AFONSO, 2009: 41).

No processo de “civilização”, as línguas autóctones eram, portanto, encaradas como línguas bastardas dos selvagens. Desprestigiadas na comunicação em público, a tais sistemas costumava ser negado um lugar oficial na sociedade. Não podiam ser ensinados, exceto para facilitar a aprendizagem do Português pelas instituições missionárias. Entre 1921 e 1954, mesmo para essa finalidade, seu uso foi impedido, sendo novamente autorizado pelo Decreto-lei n. 39666 do Estatuto dos Indígenas Portugueses das Províncias da Guiné, Moçambique e Angola<sup>49</sup>. A respeito da supremacia europeia, Feytor Pinto (2001: 66) alega: “por repudiarem a aprendizagem de línguas africanas, os portugueses preferem recorrer a intérpretes que são escravos de senhores africanos, obrigados a aprender o Português”. Além disso, “a primeira descrição portuguesa do funcionamento de uma língua africana” foi “publicada em 1697, mais de 250 anos dos primeiros contatos”. (FEYTOR PINTO, 2001: 101).

Como se verifica, no imperialismo linguístico, garante-se a hegemonia do idioma dos colonos, inferiorizando as línguas africanas. O preconceito cultural aumenta quando se reconhece a exclusividade da língua portuguesa na atribuição de cidadania. De acordo com Ferreira (1989: 29),

O colonialismo é a negação da personalidade do Outro. Ele (...) nega e reprime a cultura autóctone e obriga à cultura metropolitana. Altera os hábitos sociais, intervém na culinária, no vestuário, no sistema agrícola, no re-

---

<sup>49</sup> Cf. Diário do Governo, 20.5.1954 *apud* Feytor Pinto (2001: 89).

gime de propriedade, na habitação, no sistema jurídico, na ordem social milenar estabelecida, impõe novos padrões de cultura e substitui a língua.

Trata-se de um regime essencialmente antropofágico, que “ao instalar-se em território alheio, (...) alimenta-se de uma necessidade: a devoração do Outro. Em todos os sentidos: político, cultural, ideológico, económico, religioso, linguístico” (FERREIRA, 1989: 31). Na visão dos portugueses, contudo, essas expedições eram entendidas como uma missão de salvar, através da cidadania e da religião, os povos selvagens, como ilustram as seguintes palavras de Rosa (1973: 61):

Estamos por aqui na calma insistência de quem cumpre uma predestinação, modelando homens e povos à nossa imagem e semelhança, criando neles novos conceitos de vida, erguendo suas almas e seus corpos e, arrancando-os da comum animalidade, gravando neles a consciência de um destino transcendente (...). Esta certeza que nos empolga, de sermos obreiros da História (...) dá-nos, a nós, portugueses, um estranho complexo de grandeza.

Sucedendo a política glotofágica, inicia-se um processo de mudança linguística. O Marquês de Pombal, em 1757 e 1758, determina o ensino da língua portuguesa no Brasil em detrimento da língua geral dos índios e, mais tarde, aquela se torna o principal idioma do país. Quadro semelhante se apresenta na África, quando o general Norton de Matos, governador de Angola entre 1921 e 1924, estabelece a extinção das línguas africanas, disseminando o Português, cujo ensino se tornou obrigatório em todas as missões.

Ao longo do século XIX, o sistema dos territórios já dominados passa a diferenciar a educação oficial, destinada à população urbana europeia, da educação missionária, concedida aos povos africanos rurais. Sustentando os interesses dos colonos europeus, às missões cabia transmitir a língua portuguesa aos africanos e limitar o uso das línguas africanas ao ensino religioso. Negava-se, dessa maneira, ao Outro o direito à identidade cultural, impondo-lhe valores entendidos como mercedores de imitação. Para Afonso (2009: 46), “arreigando o seu discurso no eurocentrismo, a colonização linguística de que são vítimas os povos colonizados, desemboca na institucionalização oficial, exclusiva e definitiva da língua portuguesa”.

Em São Tomé e Príncipe, as missões portuguesas começaram desde a fase de colonização, em 1493. Já no início desse processo, o ensino religioso se destaca na tentativa de europeizar os povos locais. No entanto, conforme se esclareceu nas seções precedentes, até a independência do país, o Forro preservava-se como língua dominante, seguido pelos demais sistemas autóctones. Ali, ao mesmo tempo em que a língua portuguesa desempenha um importante papel no Império, não há uma estratégia de difusão linguística e, uma vez declarada a independência, constata-se um alto índice de analfabetismo. Para Margarido (2000), o compor-

tamento se justifica na necessidade de preservar a ignorância da população subalterna, na tentativa de limitar sua consciência crítica. Em consonância com essa perspectiva, Afonso (2009: 47) enfatiza a língua como agente específico de dominação, pois

permitiria a Portugal recuperar a sua “grandeza” consubstanciada já não numa submissão política dos demais países lusófonos, mas cultural, cabendo à variedade europeia o privilégio de continuar a gozar de estatuto hegemónico relativamente às variedades emergentes no ex-império colonial português em África. (...) A potência colonial contrariou, sensivelmente, a dinâmica cultural dos países dominados. As línguas autóctones tinham necessariamente que sofrer o efeito dessa dominação. Excluídas do ensino, proibidas de serem utilizadas, embora continuassem a ser faladas, o seu papel limitou-se, quase exclusivamente, à produção de obras orais, em alguns casos bem ricas e variadas. (...) Porém, o lugar de destaque, esse, estava reservado à língua portuguesa.

A supremacia europeia atribuiu às línguas crioulas um carácter pejorativo. Há relatos de famílias, que, por aspirarem à ascensão social, proibiam os filhos de se comunicarem nas variedades locais, usando, no ambiente familiar, formas do Português como L2 na construção da L1 das crianças nativas. A esse respeito, Mata (2010: 17) ratifica a recorrência na elite de sujeitos que não sabiam falar qualquer crioulo, mesmo nos contextos cuja resistência nacionalista é maior. Além disso, os integrantes da elite envergonhavam-se quando se viam forçados a falá-los. Nas palavras de Afonso (2009: 48), atualmente,

parece (...) reinar no espírito do filho-da-terra a ideia incutida pelo poder colonial de que o Crioulo e demais línguas nacionais são formas linguísticas bastardas, desprovidas de dignidade para serem utilizadas em espaços públicos e como tal devem circunscrever-se ao meio familiar e rural.

Em contrapartida, a língua portuguesa

Foi, efectivamente, a língua, companheira do império, embora essa companhia tenha sido partilhada com as demais línguas nacionais. (...) uma vez desmoronado o império (político), ela continua ainda a manter o seu estatuto hegemónico, (...) em regime de coexistência com as demais línguas locais.

Após a independência e as posteriores massificação do ensino e mobilidade social dela decorrente, intensifica-se o processo de mudança linguística em São Tomé e Príncipe. Havia uma língua portuguesa dominante, trazida pelo colonizador e escolhida pelos novos governantes como oficial. No entanto, como a grande maioria da população era (semi)analfabeta, o Português veicu-

lar herdou traços dos crioulos e da variedade dos Tongas, alterando em muitos aspectos o sistema linguístico alvo (cf. AFONSO, 2009: 68). Assiste-se a uma transição em que a língua de maior prestígio, a segunda língua da elite urbana, expande-se a todos os contextos de comunicação, tornando-se L1 de uma parcela considerável da população. Verifica-se, ao mesmo tempo, “um aumento gradual da erosão linguística dos crioulos, um processo tipicamente associado a contextos de línguas em contato, caracterizado pela perda de competência linguística da L1, motivada pela primazia do uso da L2” (GONÇALVES, 2016: 26).

Atualmente, as elevadas taxas de falantes (nativos) do Português caracterizam São Tomé como “um caso único na África, onde o nível de nativização das antigas línguas coloniais é geralmente baixo”<sup>50</sup> (HAGEMEIJER, no prelo). Como asseguram Gonçalves; Hagemeyer (2015: 91),

a estigmatização dos crioulos, herdada no tempo colonial, não foi devidamente ultrapassada, impedindo, em definitivo, a construção de uma identidade (...) ligada às línguas crioulas. (...) São Tomé e Príncipe é hoje a ex-colônia portuguesa onde se registra o maior número de falantes nativos do português, o que significa também que todos os crioulos autóctones de São Tomé e Príncipe estão ameaçados.

Existem, todavia, diversos estatutos para a língua portuguesa ali. O Português Europeu é a língua oficial do país, a língua-alvo dos falantes. Ao mesmo tempo, a maioria da população fala um “português santomensizado”. Tal variedade foi fortemente influenciada pelo Forro, língua materna de alguns e língua segunda de outros; e por outras línguas, como as faladas pelos contratados, atuantes nos processos históricos de aquisição da L2 e de nativização. Trata-se de “estatutos que ora remetem para uma função institucional e social que a língua cumpre junto da sociedade são-tomense no seu todo, ora para uma relação individual que com ela estabelece cada indivíduo dessa mesma sociedade” (AFONSO, 2009: 64).

Apresenta-se, pois, um processo de fossilização linguística, muito frequente em casos de contato ou aquisição de L2. Em uma situação de imersão, em que há o predomínio da norma aspirada pelo falante, a gramática manifestada alia o conhecimento da primeira língua aprendida a aspectos linguísticos mais universais. São realizações distintas do padrão, enraizadas de tal modo no discurso do falante, que dificilmente podem ser corrigidas. Ademais,

---

<sup>50</sup> Tradução de “a unique case in Africa, where the degree of nativization of former colonial languages is generally low”.

Interiorizadas, tais estruturas poderão acompanhar o falante ao longo da sua vida, vindo a ser transmitidas pelos mesmos, enquanto agentes de socialização, aos filhos ou outros elementos com os quais interagem. (...) Da convivência com as línguas crioulas o Português sofre, quer a nível da oralidade, quer na escrita, uma grande influência destas línguas, resultando numa nova variedade. (AFONSO, 2009: 69)

Por conseguinte, qualquer estudo sobre o PST deve considerar ao menos duas variedades: o Português Europeu, a língua oficial padrão, “língua materna de uma parte diminuta da população e que para a grande maioria pode ser considerada língua segunda;” e o Português local, língua veicular, “um falar de características fonológicas e morfossintáticas próprias, cujos desvios em relação à norma são significativos, em que se assiste a uma tendência cada vez maior para a contaminação do código escrito pelo código oral”<sup>51</sup> (AFONSO, 2009: 70). No último tipo, enquadra-se o Português como língua materna da maioria da população: produto do contato com as línguas crioulas e, consequentemente, repleto de desvios à norma padrão, prevista na língua oficial.

Em suma, para além das variedades crioulas, a variação intrínseca à língua portuguesa em São Tomé contribui para o multilinguismo, haja vista a convivência de: i) uma “língua materna”, de socialização da grande maioria dos falantes, o “falar são-tomense”, “variante oral do português são-tomense (...) já ela eivada de interferências, a que se vão somando outras tantas no decurso da interação social”; ii) uma “língua não materna” (alvo, oficial), que se pretende ensinar “geralmente na idade escolar”, quando o falante entraria “em contacto relativamente efectivo com o Português” em um ambiente privilegiado, a escola; e iii) uma “língua de ensino”, ministrada a crianças, a partir dos 6 anos de idade, geralmente, por professores santomenses, desprovidos do conhecimento necessário sobre a língua oficial, cujas competências linguísticas e comunicativas são limitadas (AFONSO, 2009: 71).

#### 4. CONSIDERAÇÕES FINAIS

Um breve olhar sobre a trajetória do país revela, pois, os contributos i) da primeira colonização e do ciclo do açúcar, na formação dos crioulos do Golfo da

---

<sup>51</sup> A interferência do código oral no escrito representa uma questão assaz complexa em diferentes locais onde a Língua Portuguesa é oficial. Análises sociolinguísticas defendem que, em países como Portugal e Brasil, falantes menos escolarizados e integrantes de classes sociais menos abastadas tendem a aproximar, de maneira mais explícita, os modelos de oralidade e escrita do Português. Trata-se de um comportamento que parece refletir carências educativas, como a ausência de livros e do contato com uma cultura de leitura, por exemplo.

Guiné; e ii) da segunda, durante os ciclos do café e do cacau, para a maior disseminação do Português e para a mudança linguística dela decorrente (cf. GONÇALVES, 2016: 24). Em outras palavras,

São Tomé e Príncipe sofreu uma transformação linguística e sociolinguística a partir do último quartel do século XIX que se traduz (...) na hegemonia do português em detrimento das línguas crioulas autóctones. Esta transição histórica de (...) L2 para (...) L1 reflete-se nas características que hoje constituem o português de São Tomé. (GONÇALVES; HAGEMEIJER, 2015: 92)

Há evidências históricas de que as línguas autóctones eram as mais faladas pela população até o século XX, enquanto o uso do Português se limitava “à minoria de falantes nativamente portugueses e àqueles que tinham uma relação mais próxima com o poder colonial”<sup>52</sup> (HAGEMEIJER, no prelo). Conforme o supracitado, sua ascensão no interior do país se inicia, na transição entre os séculos XVIII e XIX, com os ciclos do café e do cacau, quando a variedade europeia se torna a referência no convívio entre trabalhadores contratados e fazendeiros. Desenvolve-se, então, o chamado “Português dos Tongas”, fruto do contato entre línguas africanas dos serviços e o Português utilizado nas roças. Depois da independência, em 1975, o Português se efetiva como língua oficial da nação santomense e se amplia o acesso à educação no idioma. Visando promoção social, muitos nativos já o empregavam no seio familiar e, aos poucos, assiste-se à transição da língua de L2 para L1 de grande parte da população. De acordo com o último censo realizado no ano de 2012, 98% (170.223) da população total (173.015) declaravam-se utentes do idioma oficial. Em contrapartida, 36% (62.707), 8% (14.654), 6% (11.377) e 1% (1.753) dos habitantes alegavam falar, respectivamente, o Forro, o Caboverdiano, o Angolar e o Lung’iê.

## REFERÊNCIAS

AFONSO, H. L. *Interferências linguísticas: um contributo para o ensino da língua portuguesa em S. Tomé e Príncipe*. 2009. 184f. Dissertação (Mestrado em Língua Portuguesa) – Faculdade de Letras, Universidade de Lisboa, Lisboa, 2009.

AGOSTINHO, A. L. *Fonologia e método pedagógico do Lung’ie*. 2015. 425p. Tese (Doutorado em Filologia e Língua Portuguesa) – Faculdade de Letras, Universidade de São Paulo, São Paulo, 2015.

ALMEIDA, A. de. Da origem dos angolares habitantes da ilha de S. Tomé. *Separata das Memórias*, 1962. t. VIII.

<sup>52</sup> Tradução de “restricted to the natively Portuguese-speaking minority and to those who had a close relation with the colonial power.”

- BANDEIRA, M. *Reconstrução fonológica e lexical do protocrioulo do Golfo da Guiné*. 2017. 324f. Tese (Doutorado em Filologia e Língua Portuguesa) – Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo, 2017.
- CALDEIRA, A. Rebelião e outras formas de resistência à escravatura na ilha São Tomé (séculos XVI a XVIII). *Africana Studia*, Porto: Universidade do Porto, v. 7, p. 101-136, 2004.
- CALDEIRA, A. Variação e mudança no vocalismo do PB: perfil geo-sociolinguístico. In: CONGRESSO INTERNACIONAL DA ABRALIN, 8, 2013. Rio Grande do Norte. Comunicação.
- CALDEIRA, A. Medo e religião popular na ilha de Ano Bom. Uma aproximação histórica (séculos XVI-XIX). *Biblioteca Digital Camões*, 2008. Disponível em: <<http://cvc.instituto-camoes.pt/earcoloquiocomunicacoes>>. Acesso em: 20 maio 2017.
- CALDEIRA, A. La leyenda de lodã, o de cómo rolando, compañero del emperador Carlomagno, defendió la isla de Annobón de una invasión terrible. *Oráfrica, Revista de Oralidad Africana*, Espanha, v. 6, p. 89-114, 2010.
- CLEMENTS, J. C. Evidência para a existência de um pidgin português asiático. In: D'ANDRADE, E.; PEREIRA, D.; MOTA, M. A. (Eds.). *Crioulos de base lexical portuguesa*. Braga: Associação Portuguesa de Linguística, 2000.
- CORTESÃO, J. *Os portugueses em África*. Lisboa: Portugalia, 1968.
- CRYSTAL, D. *Language death*. Cambridge: CUP, 2000.
- FERRAZ, L. I. *The creole of São Tomé*. Johannesburgo: Witwatersrand University Press, 1979.
- FERRAZ, L. I. Portuguese creoles of West Africa and Asia. In: GILBERT, G. G. (Ed.). *Pidgin and creole languages: essays in memory of John E. Reinecke*. Honolulu: University of Hawai Press, 1987. p. 337-360.
- FERREIRA, M. Contextualização da língua portuguesa. In: FERREIRA, M. *O Discurso no percurso africano*. Lisboa: Plátano Editora, 1989. p. 311-318.
- FEYTOR PINTO, P. *Como pensamos a nossa língua e as línguas dos outros*. Lisboa: Editorial Estampa, 2001.
- GONÇALVES, R. M. G. *Propriedades de subcategorização verbal no Português de São Tomé*. 2010. 151f. Dissertação (Mestrado em Linguística) – Faculdade de Letras, Universidade de Lisboa, Lisboa, 2010.
- GONÇALVES, R. M. G. *Construções ditransitivas no Português de São Tomé*. 2016. 286f. Tese (Doutorado em Linguística) – Faculdade de Letras, Universidade de Lisboa, Lisboa, 2016.
- GONÇALVES, R. M. G.; HAGEMEIJER, T. O Português num contexto multilingue: O caso de São Tomé e Príncipe. *Revista Científica da Universidade Eduardo Mondlane, Moçambique*, v. 1, n. 1, p. 87-107, 2015.
- GRANDA, G. de. *Estudios de linguística afro-romanica*. Valladolid: Universidad de Valladolid, 1985.

GRANDA, G. de. Retenciones africanas en la fonética del crioulo português de Anón. *Revista de Filología Románica*, Madrid: Universidad Complutense Madrid, v. 4, p. 111-124, 1986.

HAGEMEIJER, T. As ilhas de Babel: a criouliização no Golfo da Guiné. *Camões, Revista de letras e culturas lusófonas*, Portugal, n. 6, p. 74-88, 1999.

HAGEMEIJER, T. *Serial verb constructions in São-Tomense*. 2000. Dissertação (Mestrado em Linguística) – Faculdade de Letras, Universidade de Lisboa, Lisboa, 2000.

HAGEMEIJER, T. Semi-lexicality and underspecification in serial verb constructions. In: CORVER, N.; van RIEMSDIJK, H. *Semi-lexical heads*. New York: De Gruyter, 2001. p. 415-451.

HAGEMEIJER, T. *Clause structure in Santomé*. 2007. Tese (Doutorado em Linguística) – Faculdade de Letras, Universidade de Lisboa, Lisboa, 2007.

HAGEMEIJER, T. Initial vowel agglutination in the Gulf of Guinea creoles. In: ABOL, E.; SMITH, N. (Eds.). *Complex processes in new languages*. Amsterdam/Filadélfia: John Benjamins Publishing Company, 2009 c. p. 29-50.

HAGEMEIJER, T. The Gulf of Guinea creoles: genetic and typological relations. *Journal of Pidgin and Creole Languages*, v. 26, n. 1, p. 111-154, 2011.

HAGEMEIJER, T. O Português em contacto em África. In: MARTINS, A. M.; CARRILHO, E. (Eds.). *Manual de Linguística Portuguesa*. Frankfurt, Leipzig: Deutsche nationalbibliothek, 2016. p. 43-67.

HAGEMEIJER, T. S. Tomé e Príncipe: labirinto e laboratório de línguas. In: SEIBERT, G. *Arquipélagos Atlânticos Crioulos. Cabo Verde e S. Tomé e Príncipe numa perspectiva comparada*, 2017. (no prelo).

HAGEMEIJER, T.; ZAMORA, A. Fa d'Ambô: past and present. *International Journal of the Sociology of Language*, v. 239, p. 193-209, 2016.

HENRIQUES, I. C. *São Tomé e Príncipe. A invenção de uma sociedade*. Lisboa: Veja, 2000.

HLIBOWICKA-WEGLARZ, B. A origem dos crioulos de base lexical portuguesa no Golfo da Guiné. *Romanica Cracoviensia*, v. 11, p. 177-185, 2012.

HOLM, J. *Pidgins and creoles*. Cambridge: Cambridge University Press, 1988.

LADHAMS, J. Article agglutination and the african contribution to the Portuguese-based creoles. *Black through White, Battlebridge*, London, UK, 2007. Disponível em: <<https://www.google.it/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&cad=rja&uact=8&ved=0ahUKewi6zKmvtpHPAhdWdxxQKHAPoA1YQFgggMAA&url=http%3A%2F>>. Acesso em: 20 jun. 2017.

LOPES, N. da S.; BAXTER, A. *Os tongas: contribuição para o estudo do Português Brasileiro*. In: SILEL, 2011, Uberlândia. Anais... Uberlândia: EDUFU, v. 2, n. 2, 2011a.

MARGARIDO, A. *A Lusofonia e os lusófonos: novos mitos portugueses*. Lisboa: Edições Universitárias Lusófonas, 2000.

- MATA, I. *Polifonias insulares: cultura e literatura de S. Tomé e Príncipe*. Lisboa: Edições Colibri, 2010.
- MAURER, P. *L'angolar. Un créole afro-portugais parlé à São Tomé*. Hamburg: Helmut Buske Verlag, 1995.
- MAURER, P. A first step towards the analysis of tone in Santomense. In: MICHAELIS, S. (Ed.). *Roots of creole structures*. Amsterdam: John Benjamins, 2008. p. 253-261.
- MAURER, P. *Principense. Grammar, texts, and vocabulary of the Afro-Portuguese creole of the island of Príncipe, Gulf of Guinea*. London: Battlebridge Publications, 2009.
- NASCIMENTO, Fabiane de Mello Vianna da Rocha Teixeira Rodrigues. *O sistema vocálico do Português de São Tomé e o comportamento das vogais médias em contexto pretônico*. 2018. 194f. Tese (Doutorado em Letras Vernáculas). Universidade Federal do Rio de Janeiro, Faculdade de Letras, Rio de Janeiro, 2018.
- NEGREIROS, A. L. de A. *História ethnographica da ilha de S. Tomé*. Lisboa: Antiga Casa Bertrand, 1895.
- POST, M. *La situación lingüística del fa d'ambô, foro hispánico (sociolingüística: lenguas en contacto)*. Amsterdam: Rodopi, 1998.
- ROSA, M. F. *O ponto e o rumo do ensino ultramarino*. Angola: Lello, 1973.
- SANTOS, C. M. A formação das estruturas fundiárias e a territorialização das tensões sociais: São Tomé, primeira metade do século XVI. *Studia*, Porto: Universidade do Porto, v. 54/55, p. 51-91, 1996.
- SCHUCHARDT, H. *Ueber das Negerportugiesische von Annobom*. Sitzungsberichte Wien, v. 116, 1888. p. 193-226.
- SEIBERT, G. Os angolares da ilha de São Tomé: náufragos, autóctones ou quilombolas? *Dossiê História Atlântica*, v. 12, n. 1/2, p. 43-64, 2004.
- SEIBERT, G. *Comrades, clients and cousins. Colonialism, socialism and democratization in São Tomé and Príncipe*. Leiden: Brill, 2006.
- SEIBERT, G. São Tomé and Príncipe. In: SPEAR, T. (Ed.). *Oxford bibliographies in african studies*. New York: Oxford University Press, 2014.
- SEIBERT, G. Substrate influence in creoles and the role of transfer in second language acquisition. *Studies in Second Language Acquisition, Cambridge*, v. 25, n. 2, p. 185-209, 2003.
- SILVEIRA, A. C. *Ditongos no Português de São Tomé e Príncipe*. 2013. 108f. Dissertação (Mestrado em Filologia e Língua Portuguesa) – Faculdade de Letras, Universidade de São Paulo, São Paulo, 2013.
- ZAMORA, A. Breve aproximación a la sociolingüística del Fa d'Ambô en Guinea Ecuatorial. *Oráfrica, Revista de Oralidad Africana*, Espanha, v. 5, p. 71-112, 2009.
- ZAMORA, A. *Gramática descriptiva del fa d'ambô*. Barcelona: CEIBA, 2010.



# 3

## CAPÍTULO

# O PERFIL MULTILÍNGUE DE MOÇAMBIQUE

**KAREN CRISTINA DA SILVA PISSURNO**

## **1. INTRODUÇÃO**

Para que se realize um estudo minimamente fiel da variedade de Português falada em Moçambique, faz-se necessário o conhecimento da realidade linguística do país, que difere, e muito, da brasileira e da europeia. A situação de contato linguístico, característica mais especificamente encontrada em algumas variedades africanas, por conta do uso simultâneo de línguas autóctones, como é o caso de Moçambique, iniciado desde sua colonização e ainda muito presente na comunidade, é um quadro fundamental, que atua invariavelmente na construção da variedade moçambicana da Língua Portuguesa.

Em termos geográficos, é importante destacar que Moçambique é um país localizado no Sudeste africano, com extensão territorial de 801.537km<sup>2</sup>, dividido em 11 províncias: Niassa, Cabo Delgado, Nampula, Zambézia, Tete, Manica, Sofala, Inhambane, Gaza, Maputo e a cidade de Maputo, capital do país (Cf. Figura 1). Em cada uma dessas províncias fala-se a língua oficial do país, o Português, e ao menos, uma língua Banto específica daquela região.



**Figura 1** Mapa de Moçambique

Fonte: Disponível em: <<https://pt.mapsofworld.com/mozambique>>

Ressalte-se que, todavia, a difusão da Língua Portuguesa pelos colonizadores portugueses nas províncias moçambicanas foi mais lenta do que no Brasil e na Índia, já que, apesar de terem chegado às terras africanas na mesma época em que as colonizavam, em 1498, Moçambique não era prioridade. Assim, o Império Português só foi anexado em 1505 e cerca de 400 anos mais tarde, em 1930, foi estabelecida uma política que tinha por objetivo educar e assimilar culturalmente os africanos, através da introdução do Português como língua de instrução escolar. Somente após a independência do país, em 1975, o Português é definido como língua oficial e passa a ser adotado (e desejado) como a maior, se não a única, fonte de comunicação geral e de transmissão da educação formal.

No entanto, mesmo depois de ser estabelecida como língua oficial do país, dados de recenseamentos (cf. CENSO, 2007; INE, 2010) e de estudos já realizados sobre a variedade moçambicana (cf. GONÇALVES, 2010; NGUNGA, 2012; TIMBANE, 2015; CHIMBUTANE, 2015; GONÇALVES; CHIMBUTANE, 2015; CAO PONSO, 2016; dentre outros) demonstram que o Português coexiste com uma grande diversidade de línguas autóctones. Esses idiomas, todos pertencentes à família Banto, são, para muitos habitantes das áreas rurais de Moçambique, especialmente aqueles acima dos 50 anos de idade, suas línguas maternas.

Os mais jovens da área rural, por sua vez, assim como os mais velhos, também têm algumas dessas línguas autóctones como maternas. Sendo assim, o idioma tido como oficial apresenta um *status* de língua estrangeira (LE)<sup>1</sup> para esses indivíduos, ou seja, uma língua utilizada em situações bastante artificiais, especialmente instrucionais, já que a língua alvo só é aprendida em contextos de educação formal, enquanto em casa os indivíduos utilizam suas línguas maternas para comunicação diária.

Por outro lado, nas áreas urbanas, a situação é similar à do uso de uma segunda língua (L2), ou seja, a exposição à língua alvo não se faz apenas em contexto escolar, mas é exigida em praticamente todos os ambientes nos quais os indivíduos estabelecem comunicação, já que, mesmo que dentro de casa eles falem sua língua materna, fora dela é necessário comunicar-se exclusivamente em outra língua, que não sua L1. No caso de Moçambique, para além do uso das línguas autóctones, a população tem contato intenso com o Português no seu dia-a-dia e nos ambientes que frequenta e, conseqüentemente, as crianças já chegam à escola com conhecimento prévio da língua. Da mesma forma, aqueles que nascem em famílias que afirmam usar somente o Português dentro de casa tendem a adquiri-lo como sua primeira língua (L1) e, igualmente, entram na escola já com conhecimento da modalidade oral do idioma. Assim, nas áreas urbanas, independentemente da idade do indivíduo, o acesso ao Português é mais fácil e constante, permitindo o contato desde muito cedo com essa língua, o que facilita sua aquisição.

Em termos gerais, portanto, pode-se dizer que o povo moçambicano, é, em sua grande maioria, pelo menos, bilíngue. O Censo 2007 corrobora essa afirmação, já que registra o Português como a língua mais utilizada por seus habitantes dentro de casa, especialmente aqueles que vivem nas áreas urbanas. Já os que vivem nas áreas rurais, por outro lado, têm na educação formal o único meio de contato com a Língua Portuguesa, fato diretamente relacionado às altas taxas de analfabetismo dos habitantes mais velhos, que continuam a utilizar as línguas Banto faladas no país.

---

<sup>1</sup> Basicamente, a diferença entre aquisição de segunda língua (L2) e de língua estrangeira (LE) está relacionada, principalmente, aos contextos nos quais as línguas são aprendidas. De acordo com Stern (1983), segunda língua está relacionada ao aprendizado e o uso de uma língua, diferente da materna, dentro de um território no qual ela exerce uma função social e/ou política. Por outro lado, uma língua estrangeira seria usada em espaços nos quais não haveria um estatuto sociopolítico que exigisse o uso da mesma. Assim, entende-se que a população rural de Moçambique não utiliza a Língua Portuguesa nos contextos que lhes são mais íntimos e familiares, pois não há exigência de uso deste idioma. Sendo assim, seu uso fica restrito às áreas nas quais o Português é a principal língua de comunicação (na cidade, nas escolas, etc.), como uma LE. Para mais informações sobre a distinção entre LE x L2, cf. Krashen (1982).

Em suma, a diversidade linguística presente nas comunidades de fala moçambicanas constitui um aspecto cultural muito relevante e tem relação direta com a questão educacional e a transmissão da Língua Portuguesa. Assim, é imperativo relatar, de forma mais detalhada, como essa diversidade de línguas é distribuída, por motivos históricos e/ou sociais, dentro da sociedade em questão, como se verá a seguir.

## 2. DADOS SÓCIO-HISTÓRICOS SOBRE AS LÍNGUAS MOÇAMBICANAS

Considerando, *à priori*, os principais perfis de falantes de Português (L2 x LE) presentes em Moçambique, estabelecem-se duas situações: de um lado, encontram-se os habitantes das zonas mais rurais, que possuem línguas *Banto* como L1 e que, teoricamente, recebem a transmissão do Português Europeu (PE) considerado padrão, como uma LE, em contextos descontínuados, formalizados apenas para os habitantes mais jovens quando em ambiente escolar; de outro lado, temos os habitantes das áreas urbanas, que têm acesso à língua em todos os seus ambientes de uso (na família, na escola e no trabalho), haja vista o valor concedido à língua oficial do país, que possibilita a ascensão social de seus usuários. Nesse caso, diversas crianças adquirem o Português como L1 (ou como uma L2 em paralelo – ou não – a uma língua autóctone), transmitido por seus pais dentro de casa e, supostamente, aprimorado na escola. A aquisição de língua para esses indivíduos, portanto, poderia atingir níveis de conhecimento parecidos com os dos falantes de PE, por exemplo, que não passam por processos semelhantes de contato com outras línguas.

Salientam-se aqui, de maneira bem breve, três fatos bastante relevantes em relação ao quadro linguístico ora apresentado: i) os pais dessas crianças, assim como alguns de seus professores, não são, necessariamente, falantes nativos de Português, o que nos leva a questionar qual “modelo” de língua lhes está sendo transmitido; ii) a importância social dada à Língua Portuguesa é salutar nessa comunidade de fala, assim como em São Tomé e Príncipe (cf. BRANDÃO; VIEIRA, 2012a; 2012b), já que seu domínio é associado ao prestígio e à ascensão econômica, desencadeando o desejo de muitos pela sua aquisição em detrimento do uso das línguas autóctones; iii) ao lado da parcela crescente de indivíduos que deseja adquirir a Língua Portuguesa e que acredita que saber (ou dizer que sabe) uma língua local pode atrapalhar suas chances de ascender socialmente, estão aqueles indivíduos, especialmente das zonas rurais, que não deixam de usar seus idiomas maternos, por seu significado afetivo e cultural.

Sendo assim, mesmo que a população reconheça o valor altamente prestigioso do Português no país, as línguas autóctones, invariavelmente presentes na rea-

lidade de todos (em maior ou menor grau), nunca chegaram a ser totalmente eliminadas e a crescente migração para a cidade faz reavivar o uso desses idiomas também dentro da cidade, nos ambientes mais diversos, sobretudo os informais. O resultado, portanto, é o uso simultâneo das diversas línguas na sociedade, que tende a selecionar os contextos em que cada uma delas tem maior valor.

Observando-se a distribuição da população de mais de 20 milhões de habitantes, que tem como língua oficial o Português, mas que, ao mesmo tempo, tem contato direto ou indireto com as mais de 20 línguas Banto (SITEO; NGUNGA, 2000 *apud* GONÇALVES, 2010, p. 25) e cinco línguas asiáticas presentes no país, confirma-se o caráter plurilíngue dessa sociedade. A tabela a seguir apresenta a relação dessas línguas, por número de falantes e províncias onde são faladas.

**Tabela 1** Línguas faladas pela população de 5 ou mais anos de idade

| Nº | LÍNGUA    | Nº DE FALANTES | %     | PROVÍNCIAS                                      |
|----|-----------|----------------|-------|---|
| 1  | Makhuwa   | 4.105.122      | 25.92 | Cabo Delgado, Nampula, Niassa, Zambézia, Sofala |
| 2  | Português | 1.828.239      | 11.54 | Todas as províncias                             |
| 3  | Changana  | 1.682.438      | 10.62 | Gaza, Maputo, Maputo City, Inhambane, Niassa    |
| 4  | Sena      | 1.314.190      | 8.30  | Manica, Sofala, Tete, Zambézia                  |
| 5  | Lomwe     | 1.202.256      | 7.59  | Napula, Niassa, Zambézia                        |
| 6  | Chuwabu   | 989.579        | 6.24  | Sofala, Zambézia                                |
| 7  | Nyanja    | 905.062        | 7.71  | Niassa, Tete, Zambézia                          |
| 8  | Ndau      | 702.455        | 4.43  | Manica, Sofala                                  |
| 9  | Tshwa     | 469.343        | 2.96  | Gaza, Inhambane, Maputo, Sofala                 |
| 10 | Nyungwe   | 457.290        | 2.88  | Manica, Tete                                    |
| 11 | Yaawo     | 340.204        | 2.14  | Cabo Delgado, Niassa                            |
| 12 | Makonde   | 268.450        | 1.69  | Cabo Delgado                                    |
| 13 | Tewe      | 255.704        | 1.61  | Manica  |
| 14 | Rhonga    | 239.333        | 1.52  | Gaza, Maputo, Maputo City, Inhambane            |

|              |                   |                   |            |                                      |
|--------------|-------------------|-------------------|------------|--------------------------------------|
| 15           | Tonga             | 203.924           | 1.38       | Gaza, Maputo, Maputo City, Inhambane |
| 16           | Copi              | 169.811           | 1.07       | Gaza, Maputo, Maputo City, Inhambane |
| 17           | Manyika           | 133.190           | 0.84       | Manica                               |
| 18           | Cibalke           | 102.778           | 0.64       | Manica                               |
| 19           | Mwani             | 77.915            | 0.49       | Cabo Delgado                         |
| 20           | Koti              | 60.780            | 0.38       | Nampula                              |
| 21           | Swahili           | 15.250            | 0.10       | Cabo Delgado                         |
| 22           | Outras            | 310.259           | 1.95       | Todas as províncias                  |
| 23           | Línguas de sinais | 7.059             | 0.05       | Todas as províncias                  |
| <b>Total</b> |                   | <b>15.833.572</b> | <b>100</b> |                                      |

Fonte: Censo de 2007<sup>2</sup>

Em síntese, podem-se destacar as seguintes línguas como concorrentes do Português na cidade de Maputo, capital do país: Changana, Tshwa, Rhonga, Tonga, Copi. Como revelam os dados extraídos do Censo 2007<sup>3</sup>, as línguas que coexistem com o Português em Maputo não são muito expressivas em se tratando de L1, sendo que a mais utilizada é a chamada Changana, falada por 31,5% da população. Vê-se, portanto, que a percentagem de habitantes que afirma usar preferencialmente o Português é bastante superior (42,9%) e aumentou em comparação com o Censo de 1997, que registrava 39,5% e assinalava o Português como a L1 de apenas 6,5% dos falantes. (cf. GONÇALVES, 2010, p. 26).

Essa diferença percentual pode ter uma de suas explicações na questão da faixa etária dos informantes do Censo, uma vez que, como a própria autora

<sup>2</sup> Tabela disponível em: <<http://www.site.letras.ufmg.br/laliafro/projeto.html>>. Acesso em: 11 de fevereiro de 2017.

<sup>3</sup> Cabe salientar que os métodos de coleta de dados para o Censo de Moçambique são recentes, tendo até o momento apenas 2 registros formais (1997, 2007). Isto nos leva a crer que é preciso cuidado ao observar os dados apresentados nas tabelas, que podem conter alguns números ainda questionáveis. Charles; Sá (2010) e Charles (2012) mencionam a necessidade de atualização das cartas geográficas que auxiliariam o processo de coleta de dados para os censos e os inquéritos feitos em Moçambique, fazendo com que seus resultados sejam mais precisos de acordo com a realidade do país.

afirma, “os índices mais elevados de conhecimento dessas línguas verificam-se entre os recenseados com mais de 25 anos” (GONÇALVES, 2010, p. 27). Isso quer dizer que, ao longo dos 10 anos que se passaram entre os recenseamentos, a formação de novas crianças, especialmente nas cidades em que o Português é a língua principal, não somente de instrução escolar, mas de uso na grande maioria dos contextos comunicativos, pode ter ajudado substancialmente a aumentar a quantidade de falantes da Língua Portuguesa na zona urbana de Moçambique como um todo, que hoje apresenta 10,7% de falantes de Português como L1, de acordo com o INE 2010, que compara a distribuição de língua materna de 1980 a 2007. Como se vê na tabela a seguir, o percentual de indivíduos que se declaram como falantes de línguas Banto como L1 vai se reduzindo proporcionalmente ao aumento de declarantes de Português como língua materna em todo o país:

**Tabela 2** Distribuição percentual da população de 5 anos ou mais anos de idade, segundo língua materna em Moçambique, em 1980, 1997 e 2007

| ANO DO CENSO | LÍNGUA BANTO | PORTUGUÊS | OUTRAS LÍNGUAS ESTRANGEIRAS | MUDO | NENHUMA | DESCONHECIDA |
|--------------|--------------|-----------|-----------------------------|------|---------|--------------|
| 1980         | 98.8         | 1.2       | –                           | –    | –       | –            |
| 1997         | 93.0         | 6.5       | 0.4                         | 0.02 | 0.06    | 1.0          |
| 2007         | 85.2         | 10.7      | 0.4                         | 0.04 | 0.01    | 3.6          |

Fonte: INE (2010) sobre Censos de 1980, 1997 e 2007

Timbane (2015), seguindo a mesma linha de pensamento, ratifica que “o número de falantes de português tende a crescer desde 1980 (5 anos após a proclamação da independência) incentivado pela educação massiva e inclusiva principalmente nas zonas urbanas” (TIMBANE, 2015, p. 295), que ocorreu tardiamente em Moçambique.

Além disso, segundo Firmino (1988),

O maior envolvimento, a todos os níveis, dos moçambicanos na acção pública (instâncias do poder, serviços públicos, comércios, etc.), o aumento das situações de comunicação em que os interlocutores não falam a mesma língua bantu [...] o aumento da população escolar atingindo níveis cada vez mais altos, são alguns dos factores que justificam a subida do número de falantes da língua portuguesa. (FIRMINO, 1988, p. 98)

Em outras palavras, a Língua Portuguesa foi, aos poucos, transformando-se na língua de comunicação básica dos moçambicanos, satisfazendo as necessidades de comunicação para obtenção de todos os tipos de serviços, especialmente nas áreas urbanas e na educação escolar, que privilegia o Português como a língua de prestígio, cujo domínio possibilita ascensão social, enquanto não dominá-la significa possibilidades de exclusão social e de preconceitos.

Timbane (2015) corrobora essas afirmações, ao atestar que

O português é uma língua moçambicana de origem europeia porque ela já satisfaz as necessidades comunicativas dos moçambicanos e já tem falantes nativos. Toda a educação formal é feita em português pelo fato de ser a língua de prestígio, por ser oficial e ser amparada pela Constituição da República de Moçambique (2004), instrumento que não dá relevância às diversas línguas bantu moçambicanas faladas pela maioria da população. (TIMBANE, 2015, p. 295)

No entanto, a população da zona rural continua tendo o Português praticamente como uma língua estrangeira, de acesso restrito ao meio escolar. Fora da escola, não há meios de difusão que permitam aos moçambicanos dessas áreas o domínio da língua de prestígio da sociedade. Menezes (2013), mencionando estudos sociolinguísticos e números do INE de 1997, com relação à área de residência em Moçambique, destaca que

os que sabem falar português nas zonas urbanas equivalem a uma percentagem de 72.4% e, nas zonas rurais, 25.4% [...], enquanto os que têm português como língua materna equivalem a 17% nas zonas urbanas e 2% nas rurais. (MENEZES, 2013, p. 38)

Já o Censo de 2007 destaca que apenas 3.5% da população rural declara ter o Português como L1. Em outras palavras, pode-se dizer que, apesar da larga expansão da Língua Portuguesa em todo o país, o número de pessoas que tem o Português como L1 é ínfimo, especialmente por conta do grande número de pessoas que utilizam as línguas Bantu nas áreas rurais.

Chimbutane (2012) confirma esse cenário, ao informar que tais dados podem

[...] ser reflexo de uma extraordinária expansão da língua portuguesa pelos meios urbanos, ao mesmo tempo que pode indicar que, apesar do crescimento da percentagem de habitantes de áreas rurais que sabem falar a língua portuguesa, estes preferem falar as línguas bantu na sua comunicação em casa. (CHIMBUTANE, 2012, p. 26)

De acordo com Gonçalves (2010), a complexidade da situação linguística do país pode ser mais bem compreendida a partir do entendimento do processo his-

tórico desde a chegada dos colonos portugueses em 1498. Apesar de terem chegado ao país na mesma época em que colonizavam o Brasil e a Índia, Moçambique não era o foco de suas atenções e, conseqüentemente, a difusão da Língua Portuguesa foi mais lenta. Sendo assim, houve uma considerável demora para a criação de políticas educacionais e, então, apenas em 1930 é estabelecida uma política, na qual o Português se torna a língua de instrução escolar, com o objetivo de educar e permitir a assimilação cultural dos africanos, deixando as línguas nacionais para uso exclusivo em contextos familiares e de instrução religiosa.

Nas palavras de Gonçalves:

Em consequência do atraso no processo de colonização de Moçambique – em que sobressai a criação tardia de uma rede escolar a nível nacional – na altura da independência, em 1975, a língua portuguesa era parte do repertório linguístico de um grupo minoritário de moçambicanos, residentes principalmente nos centros urbanos. (GONÇALVES, 2010, p. 31)

Dessa maneira, entende-se que a transmissão da Língua Portuguesa aos moçambicanos aconteceu de maneira progressiva e tardia, permitindo àqueles que tinham acesso à escola a aquisição do Português como L2 em contextos formais, enquanto os mais velhos, que não tiveram acesso à escola e são residentes das zonas rurais, permanecem utilizando suas línguas maternas Banto para comunicação natural e diária.

Após a independência do país, o Português é definido como língua oficial, tomando totalmente o lugar das línguas Banto, já que passa a ser apontada como mais eficaz para a comunicação internacional e a transmissão de conhecimentos, além de ser considerada, cada vez mais, como língua de prestígio e que permitiria ascensão social aos seus falantes. Desse modo, o Português ganha novos valores tanto para aqueles que já o dominam quanto para aqueles que desejam dominá-lo.

Advindo dessa mudança de valores, a população que frequenta a escola passou por um processo de grande crescimento, tendo, de acordo com o Censo de 2007, mais de 5 milhões de alunos. Tal número reflete a multiplicação do número de falantes da Língua Portuguesa e a importância cada vez maior facultada a ela. Gonçalves ainda assegura que

as classes mais favorecidas dos centros urbanos têm tendência a comunicar entre si exclusivamente em português (ainda que este não seja a sua L1), sendo a língua escolhida para transmitir às novas gerações. (GONÇALVES, 2010, p. 34).

Cabe ressaltar que, ao longo do período colonial, nas zonas rurais, a população continuava a se comunicar em suas línguas Banto e a transmiti-las para as

novas gerações, enquanto o Português era adquirido apenas como LE. O contato com o Português era restrito, então, à instrução formal e religiosa, e a textos escritos, de onde vinha o *input* mais estruturado possível para esses falantes. Portanto, pode-se concluir que grande parte dos falantes de Português hoje o tem como uma L2 (nas zonas urbanas) ou uma LE (nas zonas rurais), o que, conseqüentemente, gera “variabilidade das regras e traços gramaticais específicos da sua gramática” (GONÇALVES, 2010, p. 37).

Tal contexto parece sempre ter sido motivo de conflitos para os moçambicanos, especialmente durante o período colonial, pois o valor da Língua Portuguesa era tão alto, que muitas famílias proibiam suas crianças de usarem as línguas locais dentro de casa, forçando-as a utilizar apenas o Português, como declara Rosário (2015):

No tempo colonial, os pais assimilados<sup>4</sup> castigavam, algumas vezes severamente, os filhos, quando estes aprendiam as línguas africanas com as avós ou com os empregados domésticos. E as razões que apresentavam eram que, ao aprender as línguas africanas, contaminavam o seu Português, e, com isto, envergonhavam os seus pais em momentos de convívio social, por causa da “horrível pronúncia” cafricanizada, ou então porque esse conhecimento dificultava a aprendizagem e a aquisição do saber na escola, e também o desempenho no trabalho. (ROSÁRIO, 2015, p. 24)

A partir dessas crenças, as línguas autóctones foram se tornando estigmas, consideradas como línguas tribais. Aqueles que desejassem algum tipo de ascensão social e prestígio na sociedade precisavam falar Português. No entanto, devido à presença ainda muito forte dessas línguas em outros contextos, era praticamente impossível não adquirir, ainda que de forma discreta, as línguas Banto da comunidade. Assim, criou-se uma geração que, supostamente, só sabia falar Português, mas que, na verdade, nunca deixou de ter algum tipo de contato com as línguas africanas.

Assim, os jovens da minha geração, aprendendo embora as línguas maternas africanas, quer através das suas próprias mães, quer através das avós ou, no caso das famílias mais abastadas, através dos empregados domésticos, mantinham essa competência linguística mais ou menos adormecida, embora lhes fosse útil em momentos apropriados. Grande parte dos cida-

---

<sup>4</sup> Vale explicar que os “assimilados” são assim denominados por serem negros que tentaram, de muitas formas, se aproximar dos hábitos e dos costumes dos portugueses, incluindo, principalmente, o uso da Língua Portuguesa em detrimento das línguas autóctones. Seu objetivo era assimilar o status de civilização dos portugueses para obter condições similares de desenvolvimento educacional e profissional.

dãos desta geração, sobretudo os grupos das cidades mais importantes, criou a ilusão de que seus filhos não dominavam senão a língua portuguesa. (ROSÁRIO, 2015, p. 24)

Sobre os dias atuais, Chimbutane (2012) resume as tendências de uso das línguas em nível provincial, afirmando que

De modo geral, em todas as províncias, (a) as línguas *bantu* são as mais frequentemente faladas em casa, (b) o Português tende a ser cada vez mais uma língua falada com frequência no domínio familiar e doméstico.” (CHIMBUTANE, 2012, p. 29).

Isto posto, pode-se concluir que, de fato, há um plurilinguismo presente na realidade linguística de Moçambique, na qual existe predominância da Língua Portuguesa nas áreas urbanas, mas ainda há resquícios e indícios de uso dos idiomas da família Banto, principalmente entre os habitantes mais velhos da população, em ambientes domésticos e nas áreas mais rurais do país. Nas palavras de Stroud (1997),

A escolha de uma norma linguística que é externa à nação na qual a língua é falada significa que as instituições oficiais como a escola e outras arenas tais como o mercado de trabalho, acabam por exercer controlo sobre quem terá acesso à língua. Uma consequência deste facto é a criação de certo número de mercados linguísticos, mais ou menos integrados num ‘mercado’ linguístico oficial na qual os falantes podem aproximar-se da norma (europeia), e a existência simultânea e paralela de alguns mercados linguísticos não oficiais. (STROUD, 1997, p. 28).

Enquanto nas áreas mais urbanas o Português é a principal, quando não a única, língua de comunicação cotidiana, nas áreas mais rurais, por outro lado, o Português funciona como a língua de comunicação comercial, restrita a contextos de LE, convivendo com uma ou mais línguas Banto, devido ao carácter misto da população. Logo, é evidente a natureza variável da Língua Portuguesa falada nessas áreas, que convive em meio a tão intenso contato linguístico. Sendo assim, é difícil até mesmo definir a qual modelo de Português esses indivíduos, em situação tão diversificada, estão tendo acesso de fato e, ainda, quais características específicas dessa heterogeneidade estão sendo repassadas às novas gerações.

### 3. O “MODELO” DE PORTUGUÊS FALADO EM MOÇAMBIQUE

Tal contexto de extrema variabilidade implica a adaptação dos falantes rurais às variadas situações de uso do Português, dispondo, enquanto falantes de uma LE, de menos oportunidades de contato com a língua considerada padrão e

mais contato com usos não oficiais. Assim, esse cenário pode levar a variação para uma determinada direção, mais ou menos afastada do “modelo” alvo, a saber, o Português Europeu. Já os falantes urbanos, considerando também os mais escolarizados, teriam mais oportunidades de adquirir o Português padrão, à semelhança do PE.

De acordo com Chimbutane (2015), o advento da democratização do país, por volta dos anos 1980, reforçou a perspectiva de um Português padrão a ser alcançado, que excluísse totalmente qualquer característica prototípica de uma Língua Portuguesa que pudesse ser considerada propriamente moçambicana e se aproximasse, cada vez mais, da variedade europeia:

Com efeito, se bem que ainda se preconizasse o desenvolvimento de um Português moçambicano, surge uma preocupação em normalizar esse Português e até mesmo de se ter o padrão europeu como norma escolar. [...] nessa fase a preocupação central já não era a uniformização do Português falado em Moçambique [...] mas a importação de uma norma “exógena”, desconhecida pela maioria dos falantes, incluindo pela maior parte dos professores [...] a escola começou a preocupar-se mais com o ensino da gramática, passando a ser menos tolerante em relação a formas “desviantes”, incluindo aquelas que outrora eram positivamente recebidas como marcas de moçambicanidade. (CHIMBUTANE, 2015, p. 55-56)

Timbane (2014) critica a existência de um Português padrão similar ao europeu em Moçambique, afirmando que as línguas Banto ainda existem dentro e fora da cidade, sem nenhuma chance de serem totalmente extintas. Para o autor, essa situação de uso simultâneo das línguas reforça a hipótese de que existiria um Português propriamente moçambicano, “nativizado”, resultante das mudanças influenciadas pelo contato com as línguas autóctones e que geram variações diferentes do Português Europeu padrão que teria sido seu modelo de formação:

Adotamos o termo nativização [...] para designar o processo de transformação da norma-padrão europeia em PM, uma variedade que na base das LB adapta, integra na língua seus valores culturais, sua identidade, seus símbolos, seus objetos materiais de tal forma que seja sentida como pertence dos moçambicanos. (TIMBANE, 2014, p. 11)

Considerando o uso de um Português padrão, cabe ressaltar a questão social do sexo dos falantes, que parece indicar menor uso da norma de prestígio por parte das mulheres, já que estas teriam a tendência de promover o uso de línguas Banto, devido à sua baixa escolaridade e aos contextos sociais dos quais fazem parte. Ao que tudo indica, mesmo depois da independência do

país, as mulheres ainda têm menos acesso à educação formal e, conseqüentemente, à Língua Portuguesa tida como padrão. O mercado de trabalho reforçaria essa postura, haja vista seus deveres fortemente familiares e poucas oportunidades de atividade laboral fora do ambiente familiar, como esclarece Stroud (1997):

[...] embora as mulheres e os homens tenham exatamente a mesma variedade e tipo de redes sociais, os indivíduos do sexo masculino falam um Português mais padrão do que as mulheres, e estas falam melhor as línguas bantu locais que os homens. Uma explicação para isto é que as redes recebem o seu sexo de acordo com a língua; as ocupações que as mulheres têm são mais orientadas para o uso de línguas bantu locais contrariamente aos homens. (STROUD, 1997, p. 33)

No entanto, a comparação entre os Censos de 1997 (que registrava apenas 29% de mulheres que sabiam falar LP) e 2007 (apresentando o total de 41,6% na mesma categoria) demonstra estar havendo uma mudança nesse panorama, com claro crescimento no número de mulheres que declaram saber falar Português. Chimbutane (2012) discorre sobre a mudança nesse percentual e a diminuição da disparidade entre os sexos, afirmando que

Esta situação pode dever-se a uma gradual mudança de atitude em relação ao lugar da mulher na sociedade, o que pode ser consequência de diferentes políticas e acções sociais visando a equidade de gênero no país. Em termos específicos, estes dados podem indicar que a mulher tem estado a ter cada vez mais acesso à educação formal, uma das principais vias de aprendizagem da língua portuguesa em Moçambique, em especial nas áreas rurais. (CHIMBUTANE, 2012, p. 18)

Ainda assim, o autor reitera que

[...] as populações das áreas rurais e as mulheres estão numa clara situação de desvantagem em relação às populações das áreas urbanas e ao homem, respectivamente: as populações das áreas rurais e as mulheres registram índices comparativamente mais baixos de conhecimento da língua portuguesa e de alfabetização. (CHIMBUTANE, 2012, p. 43)

Em vista de toda essa desigualdade, ainda presente na sociedade, em relação ao acesso à Língua Portuguesa, seja pela localidade (rural x urbano), pela escolaridade (mais ou menos escolarizados), pela idade (indivíduos mais jovens x mais velhos) ou até mesmo o sexo (homem x mulher), a partir dos anos 90, o Ministério da Educação institucionalizou o debate sobre a introdução das línguas locais na educação formal. Assim, deu-se início às discussões sobre uma

educação bilíngue, que vem sendo apoiada por linguistas (cf. CHIMBUTANE, 2015; MENEZES, 2013; NGUNGA, 2007; LOPES, 1997; TIMBANE, 2015) que defendem a necessidade desse tipo de ensino para que, em séries mais avançadas, os alunos consigam desenvolver melhor suas habilidades, haja vista o grande fracasso escolar observado naqueles que não aprenderam Português em casa e só compreendem seus idiomas maternos. No entanto, a sua execução tem encontrado dificuldades em relação ao material didático, à grande quantidade de línguas nas quais os materiais deveriam ser produzidos e aos custos dessa produção, além da incipiente descrição das línguas locais (CHIMBUTANE, 2015).

Em poucas palavras, mesmo com toda a diversidade linguística e a tentativa de introduzir as línguas locais no ensino escolar, a Língua Portuguesa continua sendo a língua de maior prestígio e, mesmo que haja uma valorização das línguas nacionais, estas parecem estar fadadas a um simples papel integrativo, como conclui Chimbutane (2015):

O Português é assumido como um recurso que permite acesso a mercados de trabalho formais e aos dividendos socioeconômicos daí decorrentes, ao passo que as línguas locais são vistas como meros veículos de comunicação familiar ou entre membros de grupos etnolinguísticos específicos. Ou seja, no geral, as línguas locais não são associadas à geração de capital ou percebidas como recursos a explorar em mercados laborais formais. (CHIMBUTANE, 2015, p. 65)

#### **4. CONSIDERAÇÕES FINAIS**

Pode-se concluir que nem mesmo o ensino bilíngue garante melhores oportunidades de vida para os moçambicanos e que os fatores sociais são de fato relevantes no que tange à caracterização da variedade moçambicana do Português, na medida em que as variantes estigmatizadas, possivelmente decorrentes do constante contato com essas variadas línguas autóctones, não ocorrem de forma aleatória e, portanto, são regularmente estruturadas e podem ser descritas de forma a contribuir com a construção do perfil sociolinguístico desta variedade.

Tendo em vista o contexto ora apresentado, pode-se afirmar, indubitavelmente, o caráter multilíngue da sociedade moçambicana. Diante de tais circunstâncias, o estudo da variedade de Português falada em Moçambique exige que haja o cuidado especial de explorar suas especificidades. Tais propriedades fazem dela uma variedade única, demonstrando tendências semelhantes às de uma variedade ainda em processo de formação, em busca de características próprias que revelam a situação de contato linguístico, que é constante e inseparável de sua

realidade. Note-se que todo esse cenário não se deve somente aos traços sócio-políticos e culturais que se entrelaçam aos linguísticos, como também ao próprio fato de ser essa uma variedade recentemente aceita como oficial em um país independente apenas há 43 anos.

## REFERÊNCIAS

- BRANDÃO, S. F.; VIEIRA, S. R. A concordância verbal e nominal no português do Brasil e no português de São Tomé: uma abordagem sociolinguística. *Papia*, 22(1): 7-40, 2012a.
- BRANDÃO, S. F.; VIEIRA, S. R. Concordância nominal e verbal: contribuições para o debate sobre o estatuto da variação em três variedades urbanas do português. *Alfa*, v. 56, n. 3, p. 1035-1064, 2012b.
- CAO PONSO, L. O contato entre o Português e as línguas Banto em Moçambique: a alternância de atitudes sobre o estatuto social das línguas em jovens universitários. *Cadernos de Letras da UFF: línguas e culturas em contato*, n. 53, p. 121-144, 2016.
- CHARLES, A. J.; SÁ, L. A. C. M. (Cartografia Censitária de Moçambique. *III Simpósio Brasileiro de Ciências Geodésicas e Tecnologias da Geoinformação*. 2010. Recife, PE. Disponível em: <[https://www.ufpe.br/cgtg/SIMGEOIII/IIISIMGEO\\_CD/artigos/CartografiaeSIG/Cartografia/A\\_10.pdf](https://www.ufpe.br/cgtg/SIMGEOIII/IIISIMGEO_CD/artigos/CartografiaeSIG/Cartografia/A_10.pdf)> Acesso em: 12 de julho de 2016.
- CHARLES, A. J. *Proposta metodológica para cartografia censitária de Moçambique*. Dissertação de Mestrado. Universidade Federal de Pernambuco. Recife, 2012.
- CHIMBUTANE, F. S. Língua e Educação em Moçambique: Uma perspectiva Sócio-Histórica. In: GONÇALVES, P.; CHIMBUTANE, F. S. (Org.). *Multilinguismo e Multiculturalismo em Moçambique: em direção a uma coerência entre discurso e prática*. Moçambique: Alcance Editores, 2015.
- CHIMBUTANE, F. S. O uso da L1 dos alunos como recurso no processo de ensino e aprendizagem de/em Português/L2: o contexto de ensino bilíngue em Moçambique. *Revista Científica da Universidade Eduardo Mondlane*, Série: Letras e Ciências Sociais, v. 1, n. 1, p. 7-25, 2015.
- CHIMBUTANE, F. S. *Panorama linguístico de Moçambique: Análise dos dados do III Recenseamento Geral da População e Habitação de 2007*. Maputo: Instituto Nacional de Estatística, 2012.
- FIRMINO, G. Desvios à norma no português falado em Moçambique. *Actas do 4º encontro da associação portuguesa de linguística*, Lisboa, p. 97-106, 1988.
- GONÇALVES, P. Tipologia de erros do Português Oral de Maputo: um primeiro diagnóstico. In: GONÇALVES, P.; STROUD, C. (Org.). *Panorama do Português oral de Maputo*, v. II; *A construção de um Banco de "Erros"*. INDE – Cadernos de Pesquisa n. 24. Maputo: Moçambique, 1997.
- GONÇALVES, P. *A gênese do Português de Moçambique*. Lisboa: INCM, 2010.

GONÇALVES, P.; CHIMBUTANE, F. S. (Org.). *Multilinguismo e Multiculturalismo em Moçambique*: em direção a uma coerência entre discurso e prática. Moçambique: Alcance Editores, 2015.

HOLM, J. *An Introduction to pidgins and creoles*. Cambridge University Press. United Kingdom, 2000.

Instituto Nacional de Estatística. *III Recenseamento Geral da População e Habitação – 2007*: indicadores sócio-demográficos, Moçambique, Maputo, Instituto Nacional de Estatística, 2010.

JEQUE, A. *O enfraquecimento da morfologia flexional verbal (pessoa e número) no português de Moçambique*. Tese (Licenciatura) – Universidade Eduardo Mondlane, Moçambique, 1996.

JON-AND, A. *Variação, contato e mudança linguística em Moçambique e Cabo Verde. A concordância variável de número em sintagmas nominais do português*. Dissertação (Doutorado) – Stockholm University, Suécia, 2011.

JUSTINO, V. Concordância verbal em número: da descrição linguística à avaliação no contexto de ensino-aprendizagem. In: BASTOS, M.; MARQUES, J.; MONTEIRO, A.; SIOPA, C. (Org.). *Ensinar a língua portuguesa em Moçambique*: desafios, possibilidades e constrangimentos. Textos selecionados das VII Jornadas da Língua Portuguesa, p. 166-186. Porto: Porto Editora, 2015a.

KRASHEN, S. D. *Principles and Practice in Second Language Acquisition*. Oxford: Pergamon Press, 1982.

LOPES, A. J. *Política Linguística*: princípios e problemas. Maputo: Imprensa Universitária, 1997.

MENEZES, L. J. J. M. *O ensino bilíngue em Moçambique*: entre a casa e a escola. Tese (Doutorado). Salvador: UFB, 2013.

NGUNGA, A. Interferências de Línguas Moçambicanas em Português falado em Moçambique. *Revista Científica da Universidade Eduardo Mondlane*, Série: Letras e Ciências Sociais, v. 1, n. 0, p. 7-20. Moçambique, 2012.

NGUNGA, A. A intolerância linguística na escola moçambicana. *Conferência sobre educação em Moçambique*. Maputo, 2007.

PISSURNO, K. C. S. *A concordância verbal de terceira pessoa do plural na variedade moçambicana do Português*: uma abordagem sociolinguística. Dissertação (Mestrado em Língua Portuguesa) – Rio de Janeiro: Faculdade de Letras/UFRJ, 2017.

ROSÁRIO, L. A Língua Portuguesa como Factor de Desenvolvimento Nacional e Afirmação Internacional: que desafios? In: GONÇALVES, P.; CHIMBUTANE, F. S. (Org.). *Multilinguismo e Multiculturalismo em Moçambique*: Em direção a uma coerência entre discurso e prática. Moçambique: Alcance Editores, 2015.

SITOE, B.; NGUNGA, A. *Relatório do II Seminário sobre a Padronização da Ortografia de Línguas Moçambicanas*, NELIMO – Centro de Estudos das Línguas Moçambicanas, Universidade Eduardo Mondlane, Maputo, 2000.

STERN, H. *Fundamental concepts of language teaching*. Oxford. OUP, 1983.

STROUD, C. Aquisição de língua segunda. In: GONÇALVES, P.; STROUD, C. (Org.). *Panorama do Português oral de Maputo*, v. I – Objetivos e Métodos. INDE – Cadernos de Pesquisa n. 22, p. 11-45. Maputo, 1997.

TIMBANE, A. A. Desafios do ensino do português em contexto multilíngue em Moçambique. *Revista Travessias*, v. 8, n. 2. Paraná: UNIOESTE, 2015.

TIMBANE, A. A. Que português se fala em Moçambique? Uma análise sociolinguística da variedade em uso. *Revista Vocábulo*, v. VII. São Paulo, 2014.



II

## VARIÁVEIS FONÉTICO-FONOLÓGICAS



# RÓTICOS NAS VARIEDADES URBANAS SANTOMENSE E MOÇAMBICANA DO PORTUGUÊS

**SILVIA FIGUEIREDO BRANDÃO E  
ALESSANDRA DE PAULA**

## **1. INTRODUÇÃO**

Poucos são os estudos que contemplam as variedades africanas do Português, sobretudo no que respeita a aspectos fonético-fonológicos. Dentre as pesquisas já realizadas, a grande maioria versa sobre aspectos morfossintáticos, sintáticos, lexicais, entre outras áreas, como se pode verificar consultando o site<sup>1</sup> organizado pela Cátedra de Português da Universidade Eduardo Mondlane, de Moçambique. No que toca a São Tomé, entre os 53 trabalhos arrolados, apenas dois (3,7%) dizem respeito a essa área (SILVEIRA, 2013; SANTOS, 2012). Quanto à Moçambique, dos 494 trabalhos que compõem o acervo, quatro (0,8%) versam sobre questões fonológicas (EMBE, 1992; HOMO, 2002; VICENTE, 2009 e 2010).

Embora sejam marcantes alguns traços de pronúncia que distinguem as variedades santomense e moçambicana do Português Europeu, só a partir de 2015 se observa um interesse mais específico em verificar características do sistema

---

<sup>1</sup> [http://catedraportugues.uem.mz/?\\_\\_target\\_\\_=bibliografia-new](http://catedraportugues.uem.mz/?__target__=bibliografia-new)

fonológico do Português de São Tomé (PST) e do Português de Moçambique (PM). Nestes três últimos anos, foram divulgados estudos sobre questões prosódicas, o funcionamento do quadro vocálico e os róticos no PST, estes últimos também focalizados no PM.

Dentre esses temas, talvez o mais saliente seja o que concerne aos róticos, não só pela complexidade que caracteriza essa família de sons, mas também por eles apresentarem alto grau de variação no âmbito de diversas línguas e, em particular, nas variedades aqui em foco.

Assim, neste capítulo, tem-se por objetivo analisar os róticos nos contextos pré-vocálico e pós-vocálico nas variedades urbanas do PST e do PM, de acordo com os pressupostos da Teoria da Variação e Mudança (WEINREICH; LABOV; HERZOG, 1968) e nos desdobramentos da Sociolinguística Variacionista (LABOV, 1972, 1994, 2001, 2003), de forma a contribuir não só para o conhecimento dos sistemas fonológicos dessas duas variedades, mas também para apreensão de divergências e convergências entre elas e o PE, que lhes serve de norma de referência.

Desenvolve-se o tema em mais cinco seções. Na segunda, faz-se uma breve caracterização dos róticos e mencionam-se aqueles que ocorrem em variedades do Português; na terceira, indicam-se os procedimentos metodológicos que nortearam as análises. Nas seções 4 e 5, expõem-se, de forma contrastiva, os resultados referentes, respectivamente, aos contextos pré-vocálicos e pós-vocálicos no PST e no PM. Na sexta seção, apresentam-se as considerações finais.

## 2. OS RÓTICOS

Os *sons* de R, caracterizados como uma *família* (LINDAU, 1985: 167) e não propriamente como uma classe de sons, *stricto sensu*, por não apresentarem um traço comum<sup>2</sup> que os caracterize como tal, são bastante diversificados quanto à sonoridade, ao modo e ao ponto de articulação. Dessa *família* fazem parte as vibrantes alveolar e uvular, tepes ou flepes alveolar e retroflexo, aproximantes alveolar e retroflexa e as fricativas velar, uvular e glotal, que, não obstante, em diferentes línguas, são representados graficamente por *-r* ou *-rr* e se comportam de forma similar quanto à sua distribuição contextual: num sistema fonológico, em que há grupos consonantais, os róticos costumam ocorrer próximos ao núcleo silábico; em posição de coda, tendem à vocalização e ao cancelamento; e frequentemente alternam com outros róticos a depender do contexto.

---

<sup>2</sup> Ladefoged (1975) e Lindau (1978) aventam a hipótese, ainda não confirmada, de ser uma característica acústica (um terceiro formante abaixado) o elemento comum a esse conjunto de fones.

No Português Europeu, a norma de referência das variedades africanas, na posição intervocálica, a comutação de um R “fraco” por um R “forte”, como se convencionou dizer, pode acarretar distinção significativa (*caro* x *carro*). Nesse contexto, no primeiro caso, ocorre o tepe, que também está presente em coda silábica e em grupos consonantais e, no segundo, na fala atual de Lisboa, a vibrante uvular, a mesma que se encontra em início de vocábulo. Segundo Mateus; d’Andrade (2000: 11), este último segmento coocorre com outras variantes recuadas, “sobretudo a fricativa uvular sonora [ʁ] ou a surda [χ]; a vibrante alveolar [r] é comum em outros dialetos que não o aqui em consideração”<sup>3</sup> (no caso, o da região de Lisboa).

No PST e no PM não há, na fala de muitos indivíduos, o mesmo tipo de distribuição de segmentos que se verifica no Português Europeu (PE) e no Português do Brasil (PB). Em alguns casos, a neutralização entre o rótico [+ant] e o [-ant] se dá também no contexto intervocálico, em que tanto pode ocorrer um tepe quanto uma vibrante alveolar (ou uma fricativa uvular, no caso do PST), anulando distinções do tipo *caro* x *carro*. Tal variação também está presente no contexto inicial de vocábulo ([r]oça/[r]oça), em que, no PE atual, se esperaria, por exemplo, a vibrante uvular ou a vibrante alveolar (a pronúncia mais conservadora) ou, ainda, uma fricativa.

Embora seja a língua oficial em ambos os países, a situação do Português é diferente quanto à sua representatividade: enquanto, em São Tomé, é atualmente falado por 98,4% da população, predominantemente com L1, em Moçambique, apenas cerca de 10% dos habitantes o têm como L1 e 42,9%, como L2. No entanto, no que se refere à variabilidade dos róticos, há significativos pontos de contato entre as duas variedades, o que levou à formulação da hipótese da possível interferência de outras línguas faladas nessas duas áreas, que se caracterizam pelo alto grau de contato multilinguístico.

Em São Tomé e Príncipe, o Forro ou Santome, considerado a “língua nacional”, é o crioulo mais difundido na capital, tendo gozado de certo prestígio até a época da independência (1975), o que levava indivíduos de outras etnias a aprendê-lo (MAURER, 1995: 1). Antes falado por 72,4% da população, segundo o Censo de 2001, e hoje reduzido a apenas 36,3% de falantes, o Forro ainda coexiste com o Angolar, na Ilha de São Tomé; o Lung’ie (ou Principense) na Ilha do Príncipe, com o português dos Tongas e o Cabo-verdeano (crioulo de base portuguesa, nativo de Cabo Verde) e “resquícios de línguas do grupo Banto” (HAGEMEIJER, 2009), todos com um número bastante reduzido de utentes.

<sup>3</sup> “[...] namely the voiced uvular fricative [ʁ] or the voiceless one [χ]; the alveolar trill [r] is common in dialects other than the one under consideration”.

O Forro, ao que tudo indica, não tem em sua gênese nenhum rótico: quando da formação desse crioulo, os róticos dos itens lexicais do Português a ele incorporados foram reinterpretados como lateral alveolar – *faluz* (ferrugem), *ficicelo* (feiticeiro) – ou apagados – *soci* (sorte), segundo Ferraz (1979: 22-23). Só muito recentemente, itens lexicais com róticos passaram a constituir o léxico do Forro.

Quanto ao PM, em que os róticos se comportam de forma similar à do PST e o Português coexiste com cerca de 26 línguas do grupo Banto, as únicas menções à variação nessa variedade, anteriores aos estudos de Brandão; de Paula (2017a, 2017b, 2017c), encontram-se em Dias (2009, p. 405, apud TIMBANE; BERLINCK, 2012: 218) – “no PM há troca entre as consoantes vibrantes simples e múltiplas” – e em Gonçalves (2010: 42), que observa que entre os falantes de L1 Changana, a língua que predomina em Maputo,

cujo sistema fonológico possui apenas a vibrante dupla [R] e não tem uma vibrante simples [r] (Sitoe e Ngunga, 2000) pode acontecer que, em palavras do PE como *areia* ou *herói*, o [r] seja pronunciado como [R], verificando-se que, provavelmente por um fenômeno de hipercorreção, em palavras como *carro* e *morrer*, o [R] seja pronunciado como [r].

Na gramática do Changana mencionada por Gonçalves (NGUNGA; SIMBINI, 2012), entre os 22 fonemas expostos no quadro das consoantes (p. 29), aparece a “vibrante múltipla alveolar com ar pulmonar expirado. Ex.: *kurila* ‘chorar’; *rito* ‘voz’; *murimi* ‘camponês’”. Mais adiante (p. 41), diz-se que

o som representado por *r* é *vibrante*, pois na sua produção a ponta da língua bate de forma intermitente na região alveolar o que dá sensação de vibração, como nos seguintes exemplos: 20. *murimi* ‘agricultor’, *kuringa* ‘provar’, *lirimi* ‘língua,

não havendo nenhuma outra observação sobre esse segmento, que, pelo que se afirma quanto à estrutura da sílaba (p. 67), não deve ocorrer em coda:

como se vê, em Changana, quando a sílaba compreende mais do que um fonema, todos os elementos não silábicos ocorrem na margem esquerda (portanto, nunca na margem direita) e é obrigatória a presença de uma vogal que é o elemento mais proeminente da sílaba.

Se, quanto ao PM e às línguas com ele em contato, há apenas notas esparsas a respeito do comportamento dos róticos, no que tange, atualmente, ao PST, como se mencionou na introdução deste capítulo, já existem uns poucos estudos. Brandão et al. (2017, no prelo, p. 198-199) observam que

Entre 2015 e 2016, diferentes pesquisadores focalizaram os róticos no PST. No VI Encontro do Grupo de Estudos de Línguas em Contato, realizado na Universidade Federal da Bahia, Brandão; Vieira (2015) já mencionavam a variação existente nos diferentes contextos de ocorrência de R, o que também ocorreu (Brandão 2016b) durante o Encontro Anual da Associação de Crioulos de Base Lexical Portuguesa e Espanhola (ACBLPE), realizado em Cabo Verde (junho de 2016). Neste último Encontro, Bouchard (2016) informou que, na variedade emergente do PST, em variação com o tepe e com outras variantes, se produziria a fricativa uvular [ʁ] em contextos como [pʁofɛˈsoʁɐ], [bʁaziˈlejɐ], isto é, em ataque complexo e em posição intervocálica, mudança que estaria ocorrendo entre os falantes jovens da classe média “who are pride of their Santomean identity”, o que se verifica também na amostra aqui considerada, embora com uma frequência ainda baixa. Ainda em 2016, durante IX Encontro da ABECS, realizado em Brasília, na mesma sessão de trabalho, Brandão (2016c) reiterou a neutralização entre os róticos em qualquer posição e Agostinho (2016) tratou dos róticos em contexto intervocálico na fala da Ilha de Príncipe, com destaque sobretudo aos contextos também focalizados por Bouchard na sua comunicação. Acrescente-se que os róticos foram o tema da tese de Bouchard (2017).

Neste estudo, para fins de comparação e complementação da análise a ser desenvolvida no item 4, destacam-se algumas observações de Bouchard (2017), tese desenvolvida na perspectiva sociolinguística, mas caracterizada pela autora também como de natureza etnográfica.

Bouchard trata dos róticos e da expressão do pronome sujeito no PST para “discutir a emergência dessa variedade do Português por meio da compreensão de fenômenos sociais e ideológicos que explicam a variação e a mudança” (p. 294). Quanto aos róticos, seu principal objetivo é verificar, nos diferentes contextos em que eles podem ocorrer, a difusão das variantes “fortes”, sobretudo as fricativas, o que ela considera um símbolo de “santomensidade”, isto é, um traço que distinguiria o PST das demais variedades do Português, mudança que estaria sendo implementada pelos mais jovens. Para tanto, ela se vale do Programa R e de 6.720 dados da fala de 56 informantes (120 de cada um deles), distribuídos por sexo, cinco faixas etárias (12-18, 20-29, 30-39, 40-49 e 50 ou mais), nível de escolaridade e etnicidade.

Na sua amostra, ela registrou, conforme a Tabela 7.5 de sua tese (p. 254), 18,5% de fricativas (uvular, velar, glotal), 55,3% de tepes alveolares, 4,9% de vibrantes (alveolar, uvular) e 21% de cancelamento. Para fins de análise, considerou duas categorias (p. 245): R forte (fricativas e vibrantes) e R fraco (tepe e cancelamento de r).

Sua análise teve como focos principais os processos de apagamento de (-r), o uso do R forte e a emergência das fricativas róticas. Quanto ao primeiro processo,

conclui que o apagamento de R é implementado na fala dos menos escolarizados e em coda externa, sobretudo em verbos no infinitivo. Quanto aos dois outros processos, em suas conclusões afirma<sup>4</sup> que

os resultados mostraram que os santomenses usam R forte em posições não padrão, e que esse uso distinto é uma mudança em progresso claramente liderada pelas gerações mais jovens. Os resultados também indicaram que o uso do r fraco em posições não padrão é possível, refletindo uma fusão parcial entre o r fraco e o R forte. Finalmente, a fricativa emergente, que é uma alternativa à tradicional vibrante, marca a diferença entre os santomense nascidos antes e depois da independência do país (p. 297).

A análise de Bouchard não focalizou isoladamente cada contexto em que o R é suscetível de aparecer: ela levou em conta a posição no vocábulo como uma variável independente, analisada em conjunto com as demais variáveis, diferentemente da metodologia empregada nas análises desenvolvidas nos itens 4 e 5 deste estudo, em que as variantes fricativas ocorrem ainda de forma incipiente. Tal procedimento analítico, aparentemente motivado pelo principal objetivo de sua pesquisa (a emergência das fricativas), não permite observar em detalhe a distribuição e os condicionamentos que presidem ao uso das variantes pelos diferentes contextos, mas constitui uma importante contribuição para compreender por que Hagemeyer (2009: 20), entre outros, afirma que “face à situação linguística de S. Tomé e Príncipe, este será o único país da África de língua portuguesa onde a maioria da população tem actualmente o Português como primeira língua, havendo, assim, condições para a emergência de uma nova variedade”.

### 3. CORPUS E PROCEDIMENTOS METODOLÓGICOS

A análise dos róticos, realizada com apoio no Programa GOLDVARB-X, levou em conta amostras selecionadas dos *corpora* mencionados na Apresentação do livro. Todos os informantes representativos de ambas as variedades foram distribuídos por sexo, três faixas etárias (18-35 anos, 36-55 anos e 56 a 75 anos) e três níveis de escolaridade (fundamental: 5 a 8 anos; médio: 9 a 11 anos; superior). Além dessas variáveis sociais, foram consideradas, (a) para São Tomé, a variável *frequência de uso do Forro*, com o objetivo de aquilatar se o não uso do Forro ou seu uso em maior ou menor grau influenciaria a performance dos indivíduos; (b) para Moçambique, o estatuto do Português (L1 ou L2), bem como o grau de conhecimento/uso de outra(s) língua(s) falada(s) no país. Quanto às variáveis estruturais controlaram-se, a depender do contexto, o modo e o ponto de

<sup>4</sup> Esta e as demais traduções são de responsabilidade das autoras deste capítulo.

articulação das consoantes antecedente e/ou subsequente, a natureza da vogal antecedente e/ou subsequente, a classe do vocábulo, a tonicidade e o número de sílabas, bem como a classe do vocábulo.

Na seleção dos dados, levaram-se em conta os 25 minutos iniciais das entrevistas no caso do R em coda externa e interna e a duração total no que respeita ao R em onset, tendo em vista a sua baixa produtividade neste contexto.

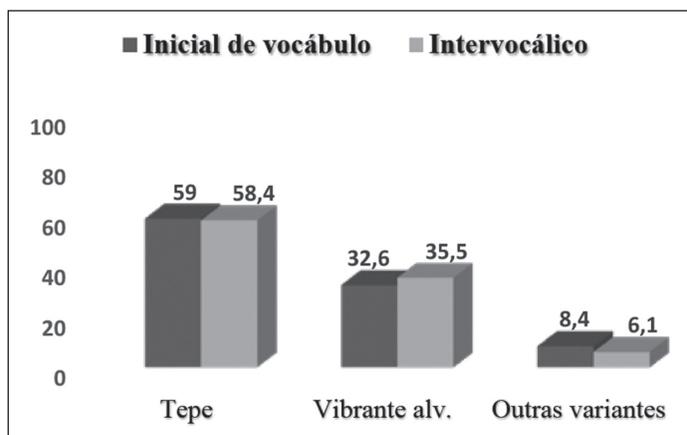
Na análise comparativa, utilizou-se um total de oito amostras, quatro por variedade e referentes (a) ao chamado “R forte” em contexto de *onset* (i) inicial de vocábulo (*roça*) e (ii) intervocálico (*carro*); e (b) ao R em coda externa (*comer*, *flor*) e interna (*carne*).

#### 4. R “FORTE” EM CONTEXTO DE ONSET

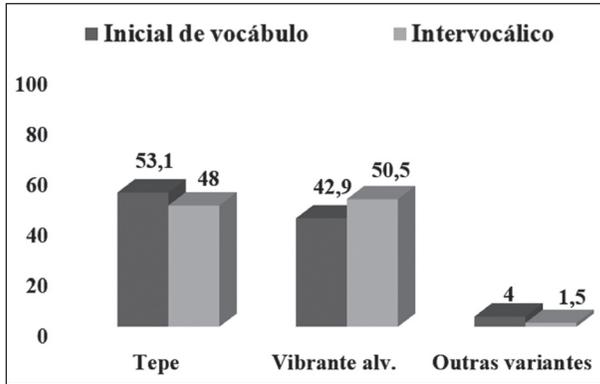
Em função dos objetivos desta análise, apresentam-se os seus principais resultados de forma contrastiva, de modo a melhor aquilatar as convergências e divergências entre as duas variedades aqui focalizadas.

Em contexto de onset inicial de vocábulo, contabilizou-se um total de 564 ocorrências de R para o PST e 752 para o PM; já no intervocálico, computou-se um total de 293 ocorrências para o PST e 529 para o PM.

Como se verifica nas Figuras 1a e 1b, em ambas as variedades, em início de vocábulo, o tepe predomina (PST: 59%; PM: 53%), enquanto no intervocálico se observa maior frequência da vibrante alveolar (35% no PST; 50,5% no PM), com diferença pouco significativa, mas que parece indicativa de que nem todos os indivíduos têm consciência da oposição fonológica como se verifica no PE e no PB. O que parece diferenciar as duas variedades é o maior percentual de variantes [ʀ] no PST, entre as quais se destaca a fricativa uvular sonora [ʀ].



**Figura 1a** Variantes de R forte em contexto pré-vocálico no PST (percentuais)

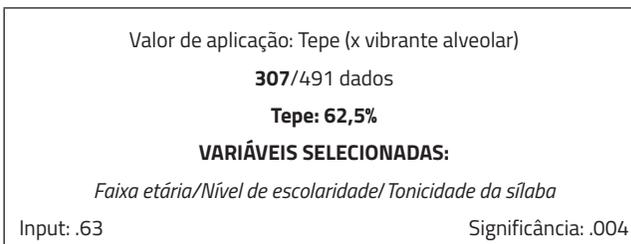


**Figura 1b<sup>5</sup>** Variantes de R forte em contexto pré-vocálico no PM (percentuais)

Tendo em vista que o tepe não é a variante esperada no PE e no PB padrão, as análises do R em *onset* terão como valor de aplicação esse fone em contraposição à vibrante alveolar, no sentido de verificar os condicionamentos que presidem à sua implementação, como se verá nos dois subitens a seguir.

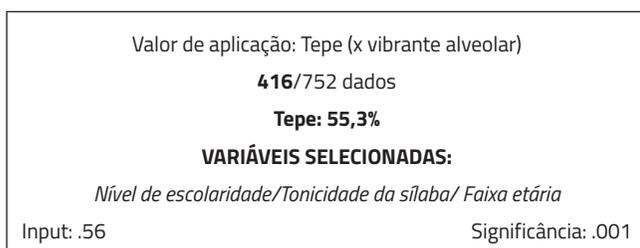
#### 4.1 Inicial de vocábulo

Em contexto inicial de vocábulo, tanto no PST quanto no PM, foram selecionadas as mesmas variáveis (cf. Figuras 2a e 2b) como salientes para a implementação do tepe, embora com diferentes relevância e *input*: .63 no PST e .56 no PM. Na primeira das citadas variedades, *faixa etária* e *nível de escolaridade* têm proeminência sobre a *tonicidade da sílaba*, enquanto, na segunda, *nível de escolaridade* e *tonicidade da sílaba* parecem exercer maior influência que a *faixa etária*. Apesar disso, em ambas as variedades, a sílaba pretônica (PST: 62,8%, P. R.: .55; PM: 61,3%, P. R.: .57) parece ser a mais suscetível ao tepe, com índices percentuais e pesos relativos bem semelhantes.



**Figura 2a** R inicial de vocábulo no PST

<sup>5</sup> Algumas das figuras e tabelas referentes à variedade do PST têm como fonte Brandão et al. (2017) e as referentes à variedade do PM, Brandão; de Paula (2017a, 2017b e 2017c).

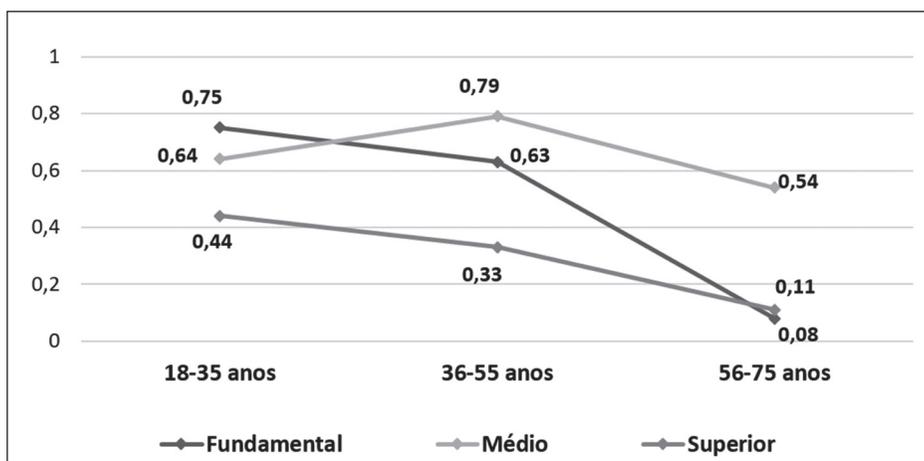


**Figura 2b** R inicial de vocábulo no PM

No que concerne à atuação da *faixa etária* e do *nível de escolaridade*, a observação dos primeiros resultados levou a procedimentos metodológicos diferentes na análise final de cada variedade. Na do PST, amalgamaram-se as duas variáveis (Figura 3), enquanto na do PM elas foram consideradas separadamente. (Figuras 4 e 5).

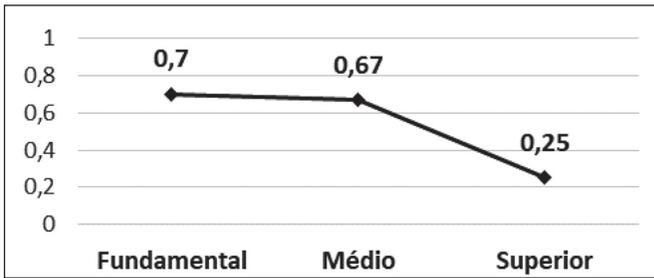
De acordo com o exposto na Figura 3, os indivíduos de nível médio de instrução têm comportamento diferenciado dos de nível fundamental e superior. Nos destas duas últimas classes, observa-se um movimento escalar, em que os mais jovens apresentam os maiores índices de uso do tepe (fundamental: P. R.: .75; superior: .44), enquanto os mais velhos detêm as menores taxas (fundamental: P. R.: .08; superior: P. R.: .11). Cabe ressaltar, ainda, que o *range* que separa os mais jovens dos mais velhos, no nível fundamental, é de .67, e no nível superior, de .33, o que parece demonstrar que os falantes deste último nível, que, em geral fizeram sua formação superior em Portugal, aproximam-se da norma europeia.

Já entre os indivíduos de nível médio, detecta-se um quadro de variação estável, em que os falantes da faixa etária intermediária apresentam os maiores índices de uso do tepe (P. R.: .79), bem como os das faixas extremas dos outros dois níveis.

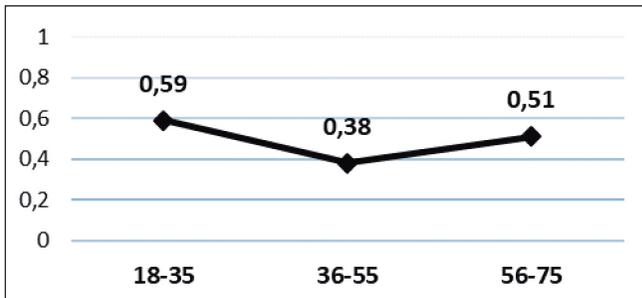


**Figura 3** Tepe em contexto inicial de vocábulo no PST – faixa etária/nível de escolaridade

No PM, detecta-se a mesma escalaridade quanto ao nível de escolaridade (Figura 4), também com os indivíduos de nível superior apresentando o menor índice (P. R.: .25) de uso do tepe, provavelmente pelo mesmo motivo indicado para o PST. Nesta variedade, a variação é estável, sendo os indivíduos de nível médio os que menos utilizam essa variante (P. R.: .38).



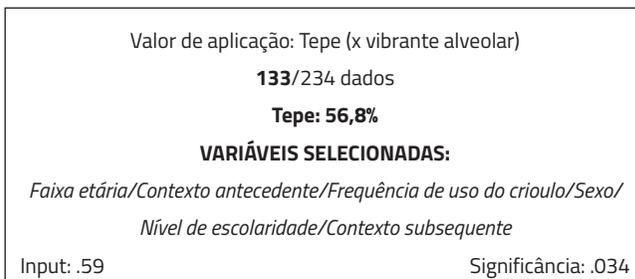
**Figura 4** Tepe em contexto inicial de vocábulo no PM – nível de escolaridade



**Figura 5** Tepe em contexto inicial de vocábulo no PM – faixa etária

## 4.2 Intervocálico

Em contexto intervocálico, aquele em que o chamado R “forte” é menos produtivo, o PST apresenta um maior número de variáveis atuando para a implementação do tepe, embora a *faixa etária* e o *nível de escolaridade* tenham sido selecionados em ambas as amostras.



**Figura 6a** R intervocálico no PST

|   |                     |
|---|---------------------|
| Valor de aplicação: Tepe (x vibrante alveolar)                  |                     |
| 254/521 dados   |                     |
| Tepe: 48,8%   |                     |
| <b>VARIÁVEIS SELECIONADAS:</b>                                  |                     |
| <i>Nível de escolaridade/Estatuto do Português/Faixa etária</i> |                     |
| Input: .48  | Significância: .000 |

**Figura 6b** R intervocálico no PM

No PM (Figura 6b), mostraram-se relevantes apenas restrições sociais, enquanto, no PST (Figura 6a), das seis variáveis selecionadas, apenas duas são de natureza estrutural. O tepe se aplica preferencialmente quando ocorrem vogais [-arr], todas com P. R. acima de .50, sendo as [-rec] – [e e i] – em contexto antecedente e a [+rec] – [a] – no subsequente, respectivamente, com P. R.: .59 e .72, as mais proeminentes, como se observa na Figura 7.

| VARIÁVEL          | FATORES              | APL./Nº             | %    | P. R.      | EXEMPLO                |
|-------------------|----------------------|---------------------|------|------------|------------------------|
| Vogal antecedente | [-arr] [+rec]        | 30/52               | 57.7 | .52        | ca[r]ego <carrego>     |
|                   | <b>[-arr] [-rec]</b> | 42/55               | 76.4 | <b>.59</b> | te[r]a <terra>         |
|                   | [+arr] [+rec]        | 54/118              | 45.8 | .44        | mo[r]er <morrer>       |
| Vogal subsequente | <b>[-arr] [+rec]</b> | 45/61               | 78.3 | <b>.72</b> | se[r]a <serra>         |
|                   | <b>[-arr] [-rec]</b> | 51/81               | 63   | <b>.64</b> | mo[r]ido <morrido>     |
|                   | [+arr] [+rec]        | 36/89               | 40.4 | .22        | co[r]upção <corrupção> |
| Input: .59        |                      | Significância: .034 |      |            |                        |

**Figura 7** Tepe em contexto intervocálico: atuação dos contextos antecedente e subsequente no PST

No que toca às variáveis sociais, apresentam-se, de início (Figuras 8a e 8b), comparativamente, aquelas que atuaram em ambas as variedades, embora, no que se refere à faixa etária, os procedimentos metodológicos tenham sido diferentes, tendo em vista que, no PST, os indivíduos mais jovens apresentaram o mesmo tipo de performance, o que determinou o amalgamento das duas faixas mais baixas.

Em ambas as amostras, os indivíduos mais velhos mostraram, claramente, menor predisposição ao uso do tepe (PST: P. R.: .22; PM: P. R.: .39) em contraposição aos mais jovens (P. R.: .63). No PM, verifica-se, com um *range* de apenas .10, uma diferença entre os mais jovens (P. R.: .64) e os da faixa intermediária (P. R.: .54), o que, de certa forma, indica que, em ambas as variedades, são os mais jovens que adotam o tepe.

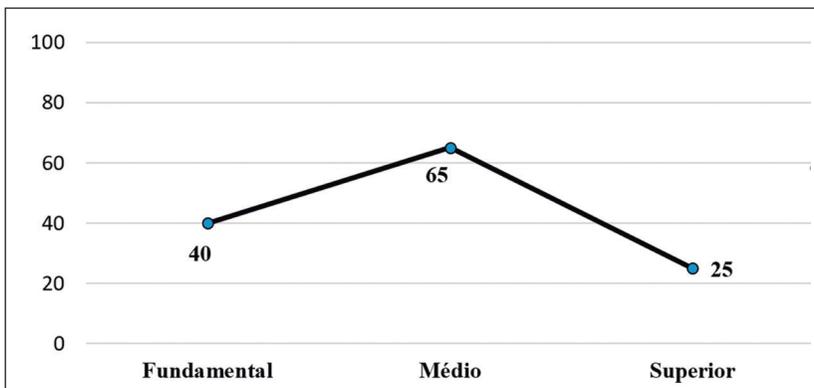
| FATORES                     | APL./Nº | %    | P. R.               |
|-----------------------------|---------|------|---------------------|
| 18-55 anos<br>(mais jovens) | 111/160 | 69,4 | <b>.63</b>          |
| 56-75 anos<br>(mais velhos) | 22/74   | 29,7 | .22                 |
| Input: .59                  |         |      | Significância: .039 |

**Figura 8a** Tepe em contexto intervocálico: atuação da variável *faixa etária* no PST

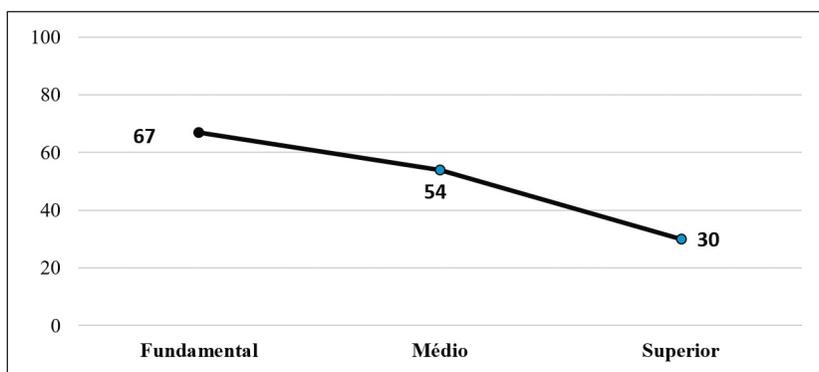
| FATORES    | APL./Nº | PERC. | P. R.               |
|------------|---------|-------|---------------------|
| 18-35 anos | 89/136  | 64,4  | <b>.64</b>          |
| 36-55 anos | 66/148  | 44,6  | <b>.54</b>          |
| 56-75 anos | 99/237  | 41,8  | .39                 |
| Input: .48 |         |       | Significância: .000 |

**Figura 8b** Tepe em contexto intervocálico: atuação da variável *faixa etária* no PM

A variável *nível de escolaridade* foi selecionada em quinto lugar na análise do PST e em primeiro, na do PM. Na variedade santomense (Tabela 9a), tanto os indivíduos de nível fundamental (P. R.: .40) quanto os de nível superior (P. R.: .25) fazem pouco uso do tepe se comparados aos de nível médio (P. R.: .65). Já na moçambicana, de forma escalar, quanto menos escolarizado o indivíduo mais ele usa o tepe (Fundamental: P. R.: .67; Superior: P. R.: .30).



**Figura 9a** Tepe em contexto intervocálico no PST



**Figura 9b** Tepe em contexto intervocálico no PM

Dentre as demais variáveis selecionadas, comentam-se, a seguir, duas outras variáveis sociais, a *frequência de uso do Forro* em relação ao PST e o *estatuto do Português*, no que tange ao PM. Ambas as variáveis buscam testar se o uso ou não de outra língua pode condicionar a escolha das variantes. Em São Tomé (Figura 10), os indivíduos que se expressam em Forro com alta ou média frequência (P. R.: .68) tendem a usar mais o tepe, da mesma forma que, em Maputo (Figura 11), aqueles que têm o Português como L2 (P. R.: .73), índices que parecem confirmar a hipótese.

|                   | APL/OCO | %    | P. R.      |
|-------------------|---------|------|------------|
| Baixa             | 72/148  | 48,6 | .39        |
| <b>Média/Alta</b> | 61/86   | 70,9 | <b>.68</b> |

**Figura 10** Tepe em contexto intervocálico: *frequência de uso do crioulo* no PST

|                 | APL/OCO | %    | P. R.      |
|-----------------|---------|------|------------|
| Língua 1        | 150/390 | 38,5 | .41        |
| <b>Língua 2</b> | 104/131 | 79,4 | <b>.73</b> |

**Figura 11** Tepe em contexto intervocálico: *estatuto do Português* no PM

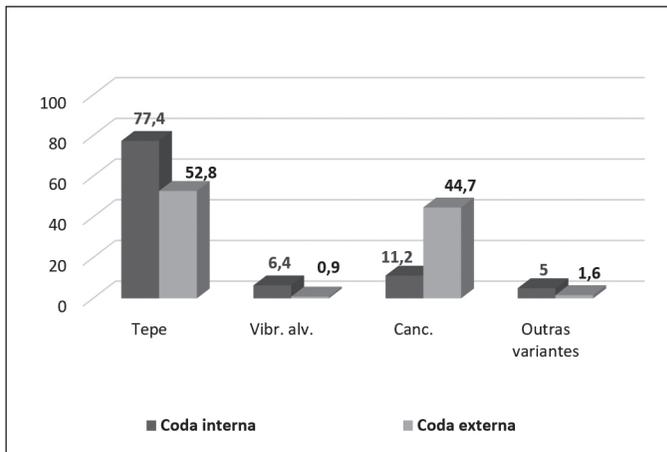
## 5. R EM CONTEXTO DE CODA

Em coda externa, levaram-se em conta 1592 dados para o PST e 2375 para o PM, observando-se, em ambas as variedades, o predomínio do tepe sobre as demais variantes (77,4% no PST e 87,3% no PM), sendo o cancelamento a segunda mais frequente (44,7% no PST e 23,5% no PM). As demais ocorrências, em

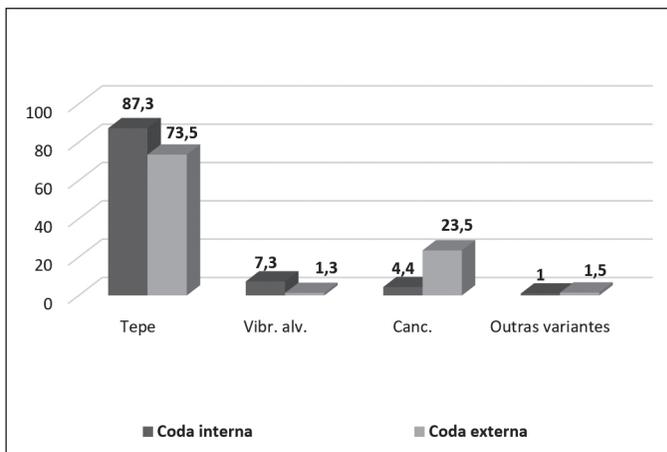
índices muito baixos (na faixa de 1%), distribuem-se pela vibrante alveolar e variantes [-ant].

Para a análise da coda interna, consideraram-se 1140 ocorrências de R no PST e 1317 no PM, dados que equivalem, em sua maioria, à realização do rótico (88,8% no PST e 95,6% no PM), uma vez que o cancelamento se mostrou muito pouco produtivo neste contexto. Este total, à semelhança do contexto de coda externa, apresenta poucos casos de variantes de R diferentes do tepe (6,4% de vibrante alveolar e 5% de outras variantes no PST; e 7,3% de vibrante alveolar e de outras variantes 1% no PM).

Nas Figuras 12a e 12b, a seguir, expõe-se a distribuição geral das variantes nos contextos de coda nas duas amostras:



**Figura 12a** Variantes de R em coda no PST (%)

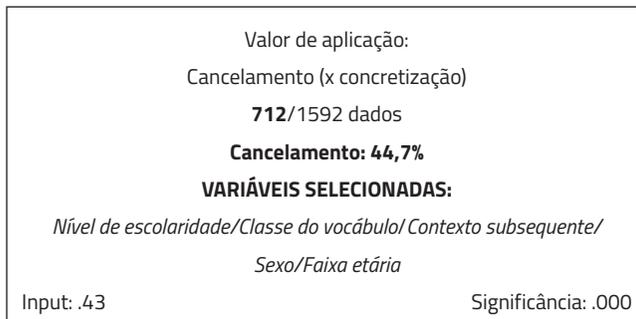


**Figura 12b** Variantes de R em coda no PM (%)

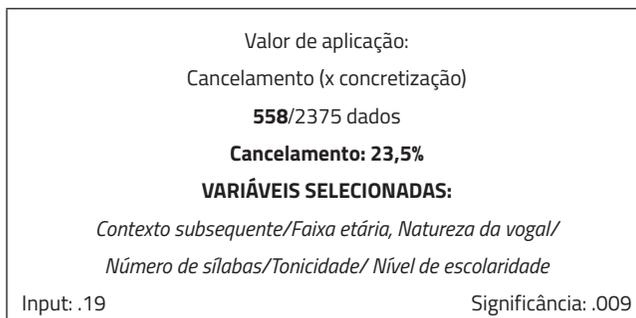
Observa-se a produtividade um pouco maior do cancelamento entre os santomenses, tanto na coda interna quanto na coda externa. Também há, no PST, percentuais um pouco maiores de outras variantes além do tepe. De qualquer forma, considerando-se a baixa frequência desses outros segmentos em ambas as variedades, na análise, o *cancelamento* foi tomado como valor de aplicação e confrontado com todas as variantes, as quais foram amalgamadas no fator *concretização*.

### 5.1 Coda externa

Como se confirma nas Figuras 13a e 13b, no PST o cancelamento está mais difundido do que no PM, sendo, não obstante, em ambos os casos, diversas as variáveis que concorrem para a implementação do processo. Foram selecionadas tanto variáveis sociais (*nível de escolaridade e faixa etária* no PST e no PM; *sexo* só no PST) quanto estruturais, dentre as quais o *contexto subsequente* em ambas as variedades.



**Figura 13a** R em coda externa no PST: índices gerais e variáveis selecionadas



**Figura 13b** R em coda externa no PM: índices gerais e variáveis selecionadas

Nas duas Figuras seguintes (14 e 15), sintetizam-se os índices obtidos para os fatores de cada variável.

| VARIÁVEL              | FATORES                    | APL./N%             | %           | P. R.      |
|-----------------------|----------------------------|---------------------|-------------|------------|
| Nível de Escolaridade | <b>Fundamental</b>         | <b>465/586</b>      | <b>79,4</b> | <b>.83</b> |
|                       | Médio                      | 149/590             | 25,3        | .29        |
|                       | Superior                   | 98/416              | 23,6        | .25        |
| Classe do vocábulo    | Nome                       | 36/207              | 17,4        | .14        |
|                       | <b>Verbo no infinitivo</b> | <b>655/1339</b>     | <b>48,9</b> | <b>.56</b> |
|                       | Outras classes             | 20/45               | 44,4        | .47        |
| Contexto subsequente  | <b>Consoante</b>           | <b>415/743</b>      | <b>55,9</b> | <b>.65</b> |
|                       | Vogal                      | 162/489             | 33,1        | .35        |
|                       | Pausa                      | 129/354             | 36,4        | .36        |
| Sexo                  | Masculino                  | 310/886             | 35          | .41        |
|                       | <b>Feminino</b>            | <b>402/706</b>      | <b>56,9</b> | <b>.60</b> |
| Faixa etária          | <b>18-35 anos</b>          | <b>354/615</b>      | <b>57,6</b> | <b>.61</b> |
|                       | 36-75 anos                 | 358/977             | 36,6        | .42        |
| Input: .43            |                            | Significância: .000 |             |            |

**Figura 14** Cancelamento em coda externa no PST: fatores condicionadores

| VARIÁVEL             | FATORES           | N%              | %           | P. R.      |
|----------------------|-------------------|-----------------|-------------|------------|
| Contexto subsequente | <b>Consoante</b>  | <b>407/1080</b> | <b>37,7</b> | <b>.69</b> |
|                      | Vogal             | 95/852          | 11,2        | .31        |
|                      | Pausa             | 55/1813         | 12,6        | .37        |
| Faixa etária         | <b>18-35 anos</b> | <b>254/844</b>  | <b>30,1</b> | <b>.59</b> |
|                      | <b>36-55 anos</b> | <b>208/846</b>  | <b>24,6</b> | <b>.52</b> |
|                      | 56-75 anos        | 96/589          | 14          | .36        |
| Natureza da vogal    | <b>[+ baixa]</b>  | <b>314/1175</b> | <b>26,7</b> | <b>.54</b> |
|                      | [-baixa]          | 244/1200        | 20,3        | .46        |
| Número de sílabas    | Uma               | 87/417          | 20,9        | .49        |
|                      | <b>Duas</b>       | <b>351/1285</b> | <b>27,3</b> | <b>.54</b> |
|                      | Três              | 103/525         | 19,6        | .45        |
|                      | Quatro ou mais    | 17/148          | 11,5        | .33        |

|                              |                    |                     |             |            |
|------------------------------|--------------------|---------------------|-------------|------------|
| <b>Tonicidade da sílaba</b>  | <b>Tônica</b>      | <b>557/2341</b>     | <b>23,8</b> | <b>.50</b> |
|                              | Postônica          | 1/34                | 2,9         | .11        |
| <b>Nível de escolaridade</b> | <b>Fundamental</b> | <b>223/903</b>      | <b>24,7</b> | <b>.50</b> |
|                              | Médio              | 157/764             | 20,5        | .44        |
|                              | Superior           | 178/708             | <b>25,1</b> | <b>.55</b> |
| Input: .19                   |                    | Significância: .009 |             |            |

**Figura 15** Cancelamento em coda externa no PM: fatores condicionadores

Levando em conta apenas os grupos de fatores comuns, verifica-se que:

- a *nível de escolaridade* atua diferentemente nas duas amostras: no PST, observa-se uma escalaridade: quanto menos escolarizado o indivíduo, mais ocorre o cancelamento; no PM, tanto os indivíduos de nível fundamental (P. R.: .50) quanto os de nível superior (P. R.: .55) deixam de concretizar o R, embora de forma neutra, o que talvez se deva ao caráter incipiente da regra (input .19) nessa variedade;
- a *faixa etária* parece atuar de forma similar, com os indivíduos mais jovens tendendo mais ao apagamento (18-35 anos, PST: P. R.: .61; PM: P. R.: .59).
- no *contexto subsequente*, a consoante é o elemento favorecedor (PST: P. R.: .65; PM: P. R.: .69), permitindo formular a hipótese de seja nessa posição que se está difundindo o apagamento.

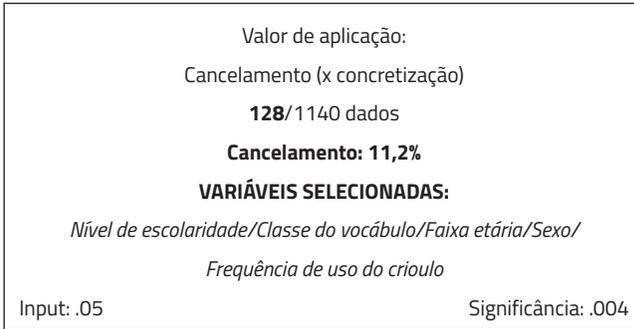
Quanto às demais variáveis, observa-se que o cancelamento é mais produtivo:

- no PST, em verbos no infinitivo (P. R.: .56), sendo as mulheres quem mais o implementa (P. R.: .60);
- no PM, as vogais [+baixas] em contexto antecedente (P. R.: .54), os dissílabos (P. R.: .54) e as sílabas tônicas (P. R.: .50), embora com índices neutros, são os fatores estruturais que favorecem o processo.

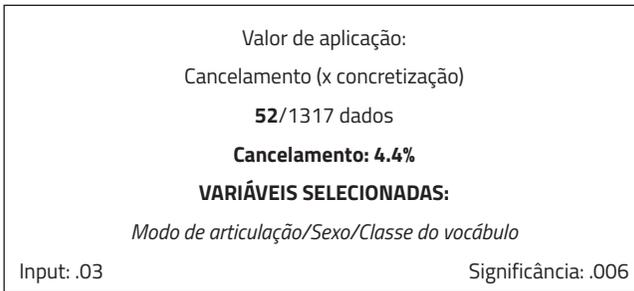
## 5.2 Coda interna

Apresentam-se as considerações sobre o cancelamento em coda interna, ressaltando que os resultados devem ser relativizados tendo em vista o baixíssimo índice de variação.

As Figuras abaixo demonstram que o cancelamento é pouco frequente nas duas variedades e parece ser diferentemente motivado em cada uma delas.



**Figura 16a** R em coda externa no PST: índices gerais e variáveis selecionadas



**Figura 16b** R em coda interna no PM: índices gerais e variáveis selecionadas

Um olhar atento sobre as amostras revela que há pouca diversidade entre os itens lexicais que as constituem, o que pode ter influenciado os resultados da análise estatística. Por exemplo, apesar da seleção da variável *classe do vocábulo* em ambas as variedades – como se verifica nas Figuras 17 e 18, abaixo –, a presença em grande quantidade das conjunções *porque* e *portanto* (itens gramaticais muito frequentes e mais passíveis de redução) na amostra de São Tomé parece estar relacionada com a relevância do fator *outras classes*, que as inclui. Em Moçambique, em que tais vocábulos não foram considerados na fase de seleção de dados, a classe dos *verbos* mostrou-se a mais saliente para o cancelamento do rótico, o que pode estar relacionado a outra das variáveis selecionadas: o modo de articulação. Das 22 ocorrências de cancelamento em verbos, 14 ocorrem quando no contexto subsequente há uma consoante [+ cont]: [s] em *conversam*, *conversar* (3 ocorrências), *exercer*, *percebo*, *percebi*, *perceber* (4 ocorrências), *proporcionar*, [ʒ] em *surgiram* e [v] em *serve*. Além disso, em Moçambique, o comportamento idiossincrático de um dos informantes do sexo masculino tornou mais salientes os índices de cancelamento nos resultados finais e pode estar relacionado com a seleção da variável *sexo*.

| VARIÁVEL                               | FATORES               | APL./N%       | %           | P. R.               |
|--|-----------------------|---------------|-------------|---------------------|
| <b>Nível de Escolaridade</b>           | <b>Fundamental</b>    | <b>88/297</b> | <b>29,6</b> | <b>.79</b>          |
|  | Médio                 | 28/488        | 5,7         | .40                 |
|  | Superior              | 12/355        | 3,4         | .34                 |
| <b>Classe do vocábulo</b>              | Nome                  | 50/725        | 6,9         | .39                 |
|  | Verbo no infinitivo   | 10/143        | 7           | .44                 |
|  | <b>Outras classes</b> | <b>68/272</b> | <b>25</b>   | <b>.77</b>          |
| <b>Faixa etária</b>                    | <b>18-35 anos</b>     | <b>65/476</b> | <b>13,7</b> | <b>.55</b>          |
|  | 36-55 anos            | 18/408        | 4,4         | .35                 |
|  | <b>56-75 anos</b>     | <b>45/256</b> | <b>17,6</b> | <b>.64</b>          |
| <b>Sexo</b>                            | Masculino             | 67/684        | 9,8         | .40                 |
|  | <b>Feminino</b>       | <b>61/456</b> | <b>13,4</b> | <b>.64</b>          |
| <b>Frequência de uso de um crioulo</b> | Zero/Baixa            | 49/543        | 9           | .45                 |
|  | Média                 | 55/551        | 10          | .51                 |
|  | <b>Alta</b>           | <b>24/46</b>  | <b>52,2</b> | <b>.80</b>          |
| Input: .05                             |                       |               |             | Significância: .004 |

**Figura 17** Cancelamento em coda interna no PST: fatores condicionadores

| VARIÁVEL                   | FATORES          | APL./N%       | %          | P. R.               |
|----------------------------|------------------|---------------|------------|---------------------|
| <b>Modo de articulação</b> | [+ contínua]     | 25/226        | 11,1       | <b>.73</b>          |
|                            | [- contínua]     | 33/1091       | 3          | .44                 |
| <b>Sexo</b>                | <b>Masculino</b> | <b>41/684</b> | <b>6</b>   | <b>.60</b>          |
|                            | Feminino         | 17/633        | 2,7        | .38                 |
| <b>Classe do vocábulo</b>  | Nome             | 34/988        | 3,4        | .46                 |
|                            | <b>Verbo</b>     | <b>22/239</b> | <b>9,2</b> | <b>.69</b>          |
|                            | Outras classes   | 2/90          | 2,2        | .37                 |
| Input: .03                 |                  |               |            | Significância: .006 |

**Figura 18** Cancelamento em coda interna no PM: fatores condicionadores

Assim, a análise multivariada do R em coda interna é aqui considerada de caráter preliminar, pretendendo-se observá-lo com mais atenção nas próximas etapas da pesquisa, especialmente com um controle mais apurado dos itens lexicais que compõem a amostra.

## 6. CONSIDERAÇÕES FINAIS

Neste capítulo, apresentaram-se, paulatina e comparativamente, os resultados das análises de (i) R “forte” em *onset* inicial de vocábulo e intervocálico e de (ii) R em coda externa e interna em duas variedades africanas do Português.

Verificou-se que, em todos os contextos, predomina o tepe, com ele concorrendo a vibrante alveolar nos casos em (i) e o cancelamento em (ii), regras condicionadas sobretudo por fatores de cunho social, salientando-se que, em contexto intervocálico, as variáveis organizadas para controlar questões relacionadas ao contato multilinguístico foram selecionadas: a *frequência de uso do Forro*, no caso de São Tomé, e o *estatuto do Português* (L1 ou L2), no caso de Moçambique.

Os resultados sugerem que, tanto no PST quanto no PM, a oposição R [+ant] / R [-ant], em contexto intervocálico não parece fazer parte da gramática muitos dos santomenses e moçambicanos, o que permite formular a hipótese de que haja um único R no quadro fonológico de ambas as variedades.

A hipótese acima formulada tem respaldo em observações de Lindau (1985). Segundo a autora, “os róticos frequentemente alternam com outros róticos”, como no Persa, em que “um /r/ vibrante tem um tepe como alofone na posição intervocálica e uma variante vibrante desvozeada na posição final de vocábulo” (p. 158). Ela afirma, ainda, que “do ponto de vista acústico, uma vibrante pode ser vista como uma série de tepes” e que “um tepe é também frequentemente uma variante de uma vibrante, particularmente na posição intervocálica” (p. 166).

A essas observações podem-se acrescentar outros argumentos: (a) em Português os pares mínimos, em contexto de oposição fonológica, têm baixíssima produtividade; (b) no PE, o tepe tem uma maior distribuição do que as demais variantes: ele ocorre em contexto intervocálico (correspondente a R [+ant]), e predomina, ainda, em ataque complexo e em coda silábica interna e externa; (c) a diferença articulatória entre um tepe e uma vibrante alveolar consiste basicamente, como afirma Lindau, entre uma batida e mais de uma batida nos alvéolos, o que pode não ser saliente para o indivíduo que adotou o Português. Além disso, não se pode esquecer que no Forro, falado em São Tomé, não há róticos (FERRAZ, 1979) e que, por exemplo, no Changana, uma das línguas de Moçambique, só há um rótico, a vibrante alveolar.

No que concerne aos resultados de Bouchard (2017), que registra a significativa ocorrência de fricativas nos diferentes contextos pré-vocálicos, cabe lembrar algumas de suas observações, no capítulo final da tese, sobre a emergência e o uso dessas variantes (p. 305):

1) [o uso das fricativas] teve um começo repentino; isso foi-me mencionado, em algumas ocasiões, por linguistas que retornaram a São Tomé e Príncipe e notaram a diferença óbvia, e por santomenses que haviam deixado as ilhas e encontrado essa pronúncia rótica quando a elas retornaram; 2) começou entre falantes urbanos e escolarizados; embora o contraste urbano / rural não tenha sido muito discutido nesta tese, considera-se que os santomenses urbanos falam melhor, e eu acho que eles têm uma taxa maior de róticos fricativos do que os santomenses rurais; 3) a fricativa provavelmente soa mais urbana, e aparentemente, “parece legal” – como foi relatado por um linguista em Príncipe, que tinha perguntado aos Principenses [sua opinião] sobre esse novo uso dos róticos.

Como já observaram Brandão et al. (2017, no prelo), o processo de mudança na realização dos róticos não se registra em plenitude nas análises por eles realizadas

mesmo tendo-se passado apenas oito anos da recolha da amostra. Sem dúvida, como já se observou em outras pesquisas sobre o PST (Brandão, 2013, entre outros), fenômenos variáveis nas variedades não europeias do Português são altamente sensíveis a condicionamentos de natureza social. Ao que tudo indica, transformações na sociedade santomense estão se processando de forma acelerada e, em consequência, motivando mudanças linguísticas.

Nesse sentido, os resultados relativos ao PST aqui expostos não só podem servir de base para estudos futuros que mostrem, com maior acuidade, a substituição do tepe e da vibrante alveolar nas posições pré-vocálicas, mas também constituem uma espécie de documento de uma fase em que se estavam apenas esboçando questões de ordem ideológica que viriam a motivar os usuários dessa variedade a encontrar um indicador de sua “santomensidade”.

Este estudo, em alguns aspectos, tem caráter preliminar. A continuidade da pesquisa poderá, portanto, determinar ajustes quanto ao que aqui se expôs. Dentre as etapas de trabalho já delineadas encontram-se: (a) ampliar o número de informantes para melhor aferição dos resultados observados, em especial no caso de São Tomé, por Bouchard (2017), que indica a tendência ao uso da fricativa uvular sonora entre indivíduos muito jovens, o que se observou em pouquíssimos dados nas amostras que serviram de base a este estudo e parece mais difundido em *onset* complexo, contexto aqui não considerado; (b) verificar a hipótese de que indivíduos com nível de escolaridade superior, tanto em São Tomé quanto em Moçambique, poderiam ter consciência do contexto de oposição fonológica, o

que também poderá ser detectado por meio de uma análise acurada das variantes mais produtivas em determinados itens lexicais; (c) realizar, no âmbito do PM, análises contrapondo falantes de Português L1 e L2.

## REFERÊNCIAS

- AGOSTINHO, Ana Livia dos Santos. *Róticos em contexto intervocálico no Português da Ilha do Príncipe: fonologia e educação*. Comunicação apresentada ao IX Encontro da ABECS. Brasília, Universidade de Brasília, 28-30 de novembro, 2016.
- BOUCHARD, Marie-Eve. *Linguistic variation and change in the Portuguese of São Tomé*. Doctoral dissertation. New York University, 2017.
- BOUCHARD, Marie-Eve. Language ideologies and use of rhotics in the Portuguese of São Tomé. Encontro Anual da ACBLPE. *Caderno de Resumos*, 18. Praia: Universidade de Cabo Verde, 2016.
- BRANDÃO, Silvia Figueiredo. Aspectos da variedade urbana do Português de São Tomé: resultados e metas de pesquisa. In: AGUILERA, Vanderci Andrade; DOIRON, Maranúbia Pereira Barbosa (Org.). *Estudos geossociolinguísticos brasileiros e europeus: uma homenagem a Michel Contini*. Cascavel-PR: EDUNIOESTE; Londrina-PR: EDUEL, 2016a. p. 67-87.
- BRANDÃO, Silvia Figueiredo. A instabilidade dos róticos no Português de São Tomé. Comunicação apresentada ao Encontro Anual da ACBLPE, 2016. *Caderno de Resumos*. Praia: Universidade de Cabo Verde, 23-25 de junho, 2016b.
- BRANDÃO, Silvia Figueiredo. *Róticos nos contextos pré e pós vocálico numa variedade africana do Português*. Comunicação apresentada ao IX Encontro da ABECS. Brasília, Universidade de Brasília, 28-30 de novembro, 2016c.
- BRANDÃO, Silvia Figueiredo; VIEIRA, Silvia Rodrigues. *Comunicação apresentada ao VI Encontro do Grupo de Estudos de Línguas em Contato*. Salvador: Universidade Federal da Bahia, 07-09 de dezembro, 2015.
- BRANDÃO, Silvia Figueiredo. Patterns of agreement within the Noun Phrase, *Journal of Portuguese Linguistics*, 12(2): 51-100, 2013.
- BRANDÃO, Silvia Figueiredo; DE PAULA, Alessandra. *Róticos em duas variedades africanas do Português*. Comunicação apresentada ao Encontro Anual da ACBLPE, 2017. Estocolmo, Universidade de Estocolmo, 13-15 de junho, 2017a.
- BRANDÃO, Silvia Figueiredo; DE PAULA, Alessandra. *Róticos no Português falado em Moçambique*. Pôster apresentado ao XXXIII Encontro Anual da Associação Portuguesa de Linguística. Évora, Universidade de Évora, 27-29 de setembro, 2017b.
- BRANDÃO, Silvia Figueiredo; DE PAULA, Alessandra. *Um estudo comparativo preliminar sobre os róticos em duas variedades africanas do português*. Comunicação apresentada ao Encontro Intermediário do GT de Sociolinguística da ANPOLL, Porto Alegre, Pontifícia Universidade Católica-RS, 22-24 de novembro, 2017c.

BRANDÃO, Silvia Figueiredo; PESSANHA, Davi Bretas dos Santos; PONTES, Stefany de Paulo. Róticos na variedade urbana do Português de São Tomé. *Papia*, 27(2): 191-213, 2017, no prelo.

DIAS, Hildizina (2009) A norma padrão e as suas mudanças linguísticas na língua portuguesa nos meios de comunicação de massas em Moçambique. In: DIAS, Hildizina (Org.). *Português moçambicano: estudos e reflexões*. Maputo: Imprensa Universitária, 2009. p. 389-420.

EMBE, Cristina. *As oclusivas orais no português em falantes nativos emakhuwa*. Tese (Licenciatura) – Universidade Pedagógica, 1992.

FERRAZ, Luiz Ivens. *The creole of São Tomé*. Johannesburg: Witwatersrand University Press, 1976.

GONÇALVES, Perpétua. *A gênese do Português de Moçambique*. Lisboa: Imprensa Nacional/Casa da Moeda, 2010.

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

HAGEMEIJER, Tjerk. As línguas de São Tomé e Príncipe. *Revista de Crioulos de Base Lexical Portuguesa e Espanhola*, Macau, 1(1): 1-27, 2009.

HOMO, Manuel. *Supressão e inserção das consoantes nasais em palavras do português por falantes nativos de tsonga: caso de Maputo*. Tese (Licenciatura) – Universidade Pedagógica, 2002.

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

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

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

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

LADEFOGED, Peter. *A course in phonetics*. New York: Hartcourt Brace Jovanovich, 1975.

LADEFOGED, Peter; MADDIESON, Ian. *The sounds of world's languages*. Oxford: Blackwell, 1996.

LINDAU, Mona. The story of /r/. In: FROMKIM, Victoria A. (Org.). *Phonetic Linguistics: essays in honor of Peter Ladefoged*. Orlando: Academic Press, 1985. p. 157-168.

LINDAU, Mona. Vowel features. *Language*, 54: 541-563, 1978.

MATEUS, Maria Helena Mira; D'ANDRADE, Ernesto. *The phonology of portuguese*. Oxford: University Press, 2000.

MAURER, Philippe. *L'Angolar: un créole afro-portugais parlé à São Tomé; notes de grammaire, textes, vocabulaires*. Hamburg: Helmut Buske Verlag, 1995.

NGUNGA, Armindo; SIMBINI, Madalena Citia. *Gramática descritiva da língua changana*. 2012. Maputo: Centro de Estudos Africanos (CEA) – UEM: 2012.

SANTOS, Eduardo Ferreira dos. Análise preliminar da palavra prosódica no português vernacular de São Tomé e Príncipe e no português falado em Angola. *Papia: Revista Brasileira de Estudos Crioulos e Similares*, 22(1): 77-89, 2012.

SILVEIRA, A. C. *Ditongos no Português de São Tomé e Príncipe*. 2013. 180fls. Dissertação (Mestrado em Filologia e Língua Portuguesa). Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, 2013.

SEIBERT, Gerhard. Crioulização em Cabo Verde e São Tomé e Príncipe: divergências históricas e identitárias. *Afro-Ásia*, 49: 41-70, 2014.

TIMBANE, Alexandre Antonio; BERLINK, Rosane de Andrade. A norma-padrão europeia e a mudança linguística na escola moçambicana. *Gragoatá* (32): 207-226, 2012.

VICENTE, Francisco. *Consciência fonológica no ensino básico em Moçambique*. Tese (Mestrado) Universidade de Lisboa, 2009.

VICENTE, Francisco. Consciência fonológica no ensino básico em Moçambique. In: GONÇALVES, Anabela; DUARTE, Inês; FREITAS, Maria João (Org.). *Avaliação da consciência linguística: aspectos fonológicos e sintáticos do Português*. Lisboa: Colibri, 2010. p. 69-93.

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

# 5

## CAPÍTULO

# AS VOGAIS MÉDIAS PRETÔNICAS NA VARIEDADE URBANA DO PORTUGUÊS DE SÃO TOMÉ

**FABIANE DE MELLO VIANNA DA ROCHA  
TEIXEIRA RODRIGUES DO NASCIMENTO**

## 1. INTRODUÇÃO

Este capítulo<sup>1</sup> descreve o vocalismo pretônico na variedade urbana do Português de São Tomé (*doravante* PST), investigando, dentre outros aspectos, a interferência dos processos de harmonização e de redução vocálicas, de fatores sociais e do crioulo mais falado na ilha, o Forro, na aplicabilidade da regra de elevação. Compara-se a variação atuante em São Tomé àquelas já descritas e observadas no Português Europeu (*doravante* PE), no Português do Brasil (*doravante* PB) e no Forro ou Santome, a fim de avaliar se o objeto de estudo em questão se insere em um *continuum* afrobrasileiro, afroeuropeu ou, ainda, se representa um sistema peculiar.

O atual panorama sociolinguístico da ilha decorre de cinco séculos de evolução, em que de língua dos colonos, o Português, paulatinamente, se transforma em língua oficial (cf. GONÇALVES; HAGEMEIJER, 2015). Ao longo de dois ciclos econômicos ocorridos na primeira e na segunda colonizações, surge um território multilíngue, em que hoje se apresentam: o Português, o Forro e três

---

<sup>1</sup> Este artigo baseia-se na tese de Nascimento (2018), de onde advieram figuras e tabelas que o compõem.

outros crioulos (o Lung'le, o Caboverdiano e o Angolar). Apesar de, atualmente, a República de São Tomé e Príncipe considerar a primeira como língua oficial, e utilizá-la como L1 ou L2, raros são os trabalhos que focalizam a variedade ali falada (cf. ARAÚJO, 2007; SILVEIRA, 2013). A maioria desses estudos concentra-se no nível sintático (cf. HAGEMEIJER, 2007, HAGEMEIJER; ALEXANDRE, 2012, GONÇALVES, 2010, entre outros), permanecendo obscuros seus aspectos fonético-fonológicos. A presente pesquisa colabora, portanto, para a descrição do quadro fonológico de uma das variedades do Português, além de abranger outras questões relacionadas ao funcionamento das línguas.

Para tanto, ressaltam-se, na seção 2, os sistemas vocálicos das variedades europeia (cf. seção 2.1) e brasileira da Língua Portuguesa (cf. seção 2.2) e do Forro (cf. seção 2.3), enfatizando a pauta pretônica. A seção 3 expõe os resultados das análises variacionistas, separadas pelas séries anterior e posterior. Após a descrição dos resultados, apresentam-se algumas considerações finais e as referências bibliográficas que nortearam a proposta.

## **2. O VOCALISMO PRETÔNICO NO PE, NO PB E NOS CRIoulos DO GOLFO DA GUINÉ**

### **2.1 No PE**

Mateus; d'Andrade (2000) descrevem o sistema vocálico do PE, afirmando que as variedades brasileira e europeia partilham sete segmentos no contexto tônico, a saber: /í, e, ε, a, ɔ, o, u/. Para os autores, na posição pretônica, há uma diferenciação evolutiva entre elas: no PB, o sistema reduz-se a cinco vogais e as médias são mais produtivas (cf. seção 2.2); já, no PE, elimina-se o valor opositivo dos mesmos sons e o quadro limita-se a dois graus de abertura que se manifestam em três vogais altas – [i, í, u] e uma vogal média [ɐ], variável com [a] tônico (p[a]go → p[ɐ]gar) (cf. MATHEUS; d' ANDRADE, 2000: 18). Depreende-se, pois, que, na última realidade, as sílabas átonas são marcadas por uma regularidade fonética maior e os quadros pré e postônicos incluem apenas quatro segmentos: /í, i, ɐ, u/<sup>2</sup>.

No plano fonético, as pronúncias dos 7 fonemas acentuados incluem o som [ɐ], cujo contraste com [a] é aparente: aquele corresponde a uma “realização alternativa de outras vogais tônicas”<sup>3</sup>, precedendo i) consoantes e glide palatais (t[ɐ]

<sup>2</sup> Compartilham dessa visão, os estudos de Gonçalves Viana (1892), Guimarães (1927), Delgado-Martins (1975), Ploae-Hanganu (1981), Mateus; Delgado-Martins (1982), Barbosa (1988), Andrade (1996), Emiliano (2009), dentre outros.

<sup>3</sup> Tradução de “an alternative realization of other stressed vowels”.

lha – <telha>; f[ɐ]cho – <fecho>; cer[ɐ]ja – <cereja>; s[ɐ]nha- <senha>; l[ɐ]i – <lei>; e ii) consoantes nasais (c[ɐ]ma; c[ɐ]na; m[ɐ]nha). Palavras oriundas de bases constituídas por [ɐ], [e] ou [ɛ] acentuados tendem a ser concretizadas com [i] na sílaba pretônica. É o que se verifica em t[ɐ]lha → t[i]lhado; l[ɐ]i → l[i]gal; s[e]lo / s[ɛ]lo → s[i]lar<sup>4</sup>. De modo semelhante, derivados de palavras formadas por [o] ou [ɔ] tônicos, como “forçar” (advindo de f[o]rça / f[ɔ]rça), se submeteriam ao alçamento da pretônica: f[u]rçar. Por conseguinte, anula-se a distinção entre vocábulos como <morar> e <murar>, posto que as vogais acentuadas [i] e [u] (v[i]vo e s[u]bo) não variam em contexto átono (respectivamente, v[i]ver, s[u]bir) (cf. MATEUS; D’ANDRADE, 2000: 19).

Tendo em vista a estrutura da sílaba e suas consequências no nível fonético, Mateus; d’Andrade (2000: 58) defendem a presença obrigatória de ataque ou rima, admitindo o não preenchimento de apenas uma dessas posições. Em decorrência da falta de um constituinte, algumas vogais podem não ser pronunciadas no PE, seja i) pela ausência de vogal subjacente, isto é, pelo fato de o núcleo não estar associado ao segmento ([ʃ]cola; [ps]icologia; ca[pt]ar; [pn]eu; ri[tm]o; a[mn]ésia; a[bs]urdo; a[ft]a); ou ii) pelo apagamento de uma vogal subjacente ([pk]eno; [ds]ifrar; [tlfɔn]; [dʃprvniɾ]; [dʃprʃtiziár]). Formam-se, assim, no nível fonético, grupos consonânticos constituídos da união de: i) duas plosivas; ii) uma plosiva e uma nasal; iii) duas nasais em posição inicial ou medial de palavra; iv) uma plosiva e uma fricativa, ou vice-versa.

Os casos de apagamento de uma vogal subjacente envolvem o encontro de até três consoantes na estrutura superficial, violando o princípio da sonoridade. Deve-se pressupor a existência de vogais na estrutura subjacente, resgatadas em palavras morfológicamente relacionadas (d[e]vo, [dvér] e [dvdór]). Ademais, o ataque vazio parece bloquear as pronúncias [i] e [u], sendo muito mais prováveis ocorrências como: a) [ɛ]lvira (“Elvira”); b) [e]/[i]lefante (“elefante”), [e]/[i]rmida (“ermida”), [e]/[i]vidente (“evidente”); e c) [o]/[ɔ]rnar (“ornar”), [o]/[ɔ]spital (“hospital”), [o]/[ɔ]lhar (“olhar”). Em a), a vogal aberta [ɛ] seria proveniente da lateral em coda. Em b), os timbres médio-alto [e] e alto [i] variam; e, em c), as vogais /o/ e /ɔ/ subjacentes não se elevam, sendo mais produtivas as pronúncias médias.

## 2.2 No PB

Câmara Jr. (1970) propõe que o número de fonemas vocálicos do PB varia de acordo com a posição do segmento quanto ao acento, atingindo seu limite

<sup>4</sup> Nesses casos, frequentemente o segmento átono é apagado no discurso coloquial, sendo comum pronúncias como [slár] para /se’laR/.

máximo de distinção na matriz tônica. Postula a existência, de sete unidades distintivas acentuadas que se reduzem, por neutralização, a cinco, quatro e três fonemas em ambientes pretônico, postônico não final e final, respectivamente.

Nas sílabas pretônicas, a neutralização efetiva-se entre os segmentos médios abertos e fechados, tanto na série anterior quanto na posterior. Tal fato se comprova em palavras como “pesar” e “morar”, cujas pronúncias abertas ou fechadas das vogais (p[ɛ]sar ou p[e]sar e m[ɔ]rar ou m[o]rar, respectivamente) não implicam mudança de sentido, como ocorre em contexto tônico: são apenas variantes dos chamados arqui-fonemas /e/ e /o/. A maior frequência de uma ou de outra variante na fala das principais áreas geográficas do Brasil parece confirmar a distinção básica entre os dialetos do Norte e do Sul do país (cf. NASCENTES, 1953). A Tabela 1 expõe dados do Projeto ALIB – Atlas Linguístico do Brasil – a respeito de 25 capitais de estado<sup>5</sup> e evidencia o predomínio de uma das pronúncias médias nos dialetos (cf. MOTA; CARDOSO, 2006). Nas capitais da região Nordeste, cerca de 60% das médias pretônicas são abertas. Aquelas localizadas no Norte e no Centro-Oeste ocupariam uma posição intermediária, “com maior frequência das médias abertas (36%, no Norte, e 27,5%, no Centro-Oeste) do que nas (...) do Sul e Sudeste (percentuais abaixo de 10%), mas com menor frequência do que nas do Nordeste”.

**Tabela 1** Vogais médias pretônicas: resultados das capitais em Mota; Cardoso (2006: 5-6)

| REGIÕES        | VOGAIS MÉDIAS PRETÔNICAS |              |          |              |
|----------------|--------------------------|--------------|----------|--------------|
|                | ABERTAS                  |              | FECHADAS |              |
| Norte          | 1.136                    | 36%          | 2.008    | <b>64%</b>   |
| Nordeste       | 1.776                    | <b>60,7%</b> | 1.149    | <b>39,3%</b> |
| Sul            | 176                      | 9,5%         | 1.683    | <b>90,5%</b> |
| Sudeste        | 210                      | 8,6%         | 2.241    | <b>91,4%</b> |
| Centro – Oeste | 739                      | 27,5%        | 1.953    | <b>72,5%</b> |

Os segmentos [e], [ɛ], [o] e [ɔ] podem, ainda, coatuar com os altos [i] e [u]. A última possibilidade, classificada como alçamento ou alteamento e facilmente verificável em todos os dialetos do PB, culmina em uma neutralização esporádica entre vogais médias e altas, conhecida como debordamento. Câmara Jr.

<sup>5</sup> Não foram analisadas a fala de Palmas, capital do estado de Tocantins, recentemente criada, bem como a de Brasília, pelo mesmo motivo.

(1970,1977), Bisol (1981, 2003), Callou & Leite (2005), entre outros renomados fonólogos e foneticistas, registram ocorrências de alteamento, principalmente, em vocábulos que apresentam uma vogal alta tônica sucedendo a pretônica, em ambas as séries. A harmonia vocálica não se restringe, entretanto, a ambientes em que atua a vogal alta: por assimilação, a presença de uma vogal média alta ou média baixa e mesmo baixa na palavra favorecerá, respectivamente, a atuação dos processos de manutenção e de abertura de timbre.

Há, enfim, casos em que “as vogais se tornam articulatoriamente mais próximas dos segmentos consonantais adjacentes”, compondo um contexto fonológico em que “não há, na palavra, vogal alta que possa desencadear o (...) alçamento” (CARMO, 2009: 24). Dito de outra maneira, os traços distintivos das consoantes precedente e/ou subsequente à vogal alvo, muitas vezes, inibem ou condicionam um processo. Resultados de Bisol, divulgados em Bortoni; Gomes; Malvar (1992: 14-15), indicam a maior incidência da elevação,

na presença de consoantes que tenham articulação alta: palatais e velares. Já a vogal em início de palavra, sem travamento silábico, mostrou ser ambiente refratário à elevação no dialeto carioca (cf. CALLOU; LEITE, 1986b). No caso do /o/, a elevação nesse ambiente praticamente não existe. No corpus de Callou; Leite, há registro de elevação nessa posição: [u]rigem. (...). Ainda, segundo Bisol, a consoante alveolar deverá desfavorecer a elevação.

### 2.3 No Forro

Entre as poucas pesquisas fonético-fonológicas sobre as línguas crioulas faladas nas Ilhas de São Tomé e Príncipe já divulgadas, priorizam-se as descrições de Ferraz (1979, 1987) e de Hagemeyer (2009c) sobre o Forro, crioulo majoritário na região. Os estudos de Ferraz representam análises gerais e pioneiras que apontam aspectos fonológicos, morfológicos, sintáticos e lexicais desse (cf. FERRAZ, 1979) e de outros crioulos de base portuguesa, encontrados na Ásia e no Oeste da África (cf. FERRAZ, 1987). Aprofundando algumas observações expressas em tais abordagens, Hagemeyer (2009 c) explica o processo de aglutinação das vogais iniciais nos crioulos do Golfo da Guiné e interpreta a harmonia vocálica como uma das possíveis motivações para o fenômeno.

Segundo Ferraz (1979: 29), o Forro apresenta uma série de mecanismos, (morfológicos e fonológicos), que viabilizam a incorporação de palavras da língua de superstrato, o Português Europeu. Como os demais crioulos do Golfo da Guiné, seu sistema vocálico se constitui de 7 segmentos orais e se submete ao processo de harmonia, fenômeno recorrente nas línguas em geral, partilhado pelo Português e por outros idiomas dos grupos Banto e Kwa (cf. FERRAZ, 1979: 49). Mais precisamente, nesses crioulos, constatam-se “exemplos de um tipo de har-

monia vocálica que é a preferência em vez da assimilação completa”<sup>6</sup> (FERRAZ, 1979: 19) (cf. Quadro 1). Trata-se de uma característica fonológica essencial, baseada em aspectos fonêmicos e morfofonêmicos: uma tendência, nem sempre sistemática, “à ocorrência da mesma vogal em duas sílabas consecutivas dentro de um morfema”<sup>7</sup> (FERRAZ, 1979: 25).

**Quadro 1** Palavras do Forro, em que se verifica o processo de harmonia vocálica, selecionadas por Ferraz (1979)

| PALAVRA DE ORIGEM<br>(PORTUGUÊS) | PALAVRA INCORPORADA<br>(FORRO) | TRADUÇÃO  |
|----------------------------------|--------------------------------|-----------|
| /ˈkɔrdə/                         | [ˈkɔdɔ]                        | “Corda”   |
| /ˈdɔr/                           | [ˈdolo]                        | “Dor”     |
| /izɐˈbɛlɔ/                       | [zɛˈbɛ]                        | “Isabel”  |
| /ˈlɑrɡu/                         | [ˈlalugu]                      | “Largo”   |
| /ˈmɛl/                           | [ˈmɛle]                        | “Mel”     |
| /ˈɔdyu/                          | [ˈɔjɔ]                         | “Ódio”    |
| /pəˈdir/                         | [piˈji]                        | “Pedir”   |
| /ˈpoku/                          | [ˈpoko]                        | “Pouco”   |
| /prəˈsizu/                       | [pliʃizu]                      | “Preciso” |
| /sɐˈbɛr/                         | [sɛˈbɛ]                        | “Saber”   |
| /ˈvɛrdə/                         | [ˈvede]                        | “Verde”   |
| /vɛrˈdadə/                       | [vɛˈde]                        | “Verdade” |

O acento da palavra incorporada ao crioulo corresponde àquele da forma original do Português. Nesse sentido, respeitando certas regras habituais de conversão de fonemas, a vogal do Forro, determinante para a harmonia, seria absorvida da língua de superstrato, tendendo a aproximar-se articulatoriamente dela.

<sup>6</sup> Tradução de “In ST there are also examples of a type of vowel harmony which is preference rather than complete assimilation. Usually the preference is not completely systematic”.

<sup>7</sup> Tradução de “vowel harmony consists in a tendency for the same vowel to occur in two consecutive syllables within a morpheme”.

Hagemeijer (2009c) avalia a recorrência da aglutinação de vogais iniciais em itens etimologicamente iniciados por consoantes na língua de superstrato. Estudos mais recentes revelam que os crioulos do Golfo da Guiné partilham não só a base lexicadora, mas também, derivam da expansão no tempo e no espaço de uma protolíngua. A aglutinação de vogais iniciais no protocrioulo seria determinada pelo gênero em Português e por um reduzido sistema-ATR, aparentemente de origem edóide. O fenômeno é, assim, motivado por questões morfológicas de gênero e fonológicas de harmonia vocálica (cf. HAGEMEIJER, 2009c: 38).

Reiterando o exposto em Ferraz (1979), Hagemeijer (2009) afirma que os crioulos do Golfo da Guiné utilizam o mesmo inventário vocálico do Português padrão com 7 vogais orais [a e ε i o ɔ u]. Aquela tendência à mesma vogal ocorrer em duas sílabas consecutivas dentro de um morfema (cf. FERRAZ, 1979: 25) deve, todavia, ser interpretada como um caso de assimilação, pois a harmonia vocálica nos crioulos em foco é um fenômeno mais restrito ao domínio das vogais médias (cf. HAGEMEIJER, 2009c: 37). Não existem muitos trabalhos sobre as restrições fonológicas envolvidas na aglutinação de vogais iniciais, mas, no Forro são recorrentes casos como os listados no Quadro 2. Essas palavras dissilábicas indicam que, nesse sistema, i) as vogais baixa e altas não determinam a qualidade dos sons adjacentes; e ii) se manifestam restrições de coocorrência motivadas pelas vogais médias. Ao menos em itens dissilábicos, verifica-se, então, uma regra de harmonia consistente que restringe a coocorrência de segmentos médios abertos e médios fechados (cf. HAGEMEIJER, 2009c: 37).

**Quadro 2** Palavras dissilábicas em Forro, extraídas de Hagemeijer (2009c: 37)

| V1 | V2    | i    | u                   | e     | ε    | o      | ɔ    | a |
|----|-------|------|---------------------|-------|------|--------|------|---|
| i  | Ligi  | Migu | izê (esteira, cama) | Mile  | libo | Jinklo | Mina |   |
| u  | Buli  | Mulu | Ubwê (boi)          | kume  | –    | –      | Uswa |   |
| e  | Sêji  | Dêsu | Vêndê               | –     | –    | –      | Zema |   |
| ε  | Peli  | Petu | –                   | vede  | –    | Tebo   | Bega |   |
| o  | Sôtxi | Wôdu | Ômê                 | –     | pôvô | –      | Lopa |   |
| ɔ  | Doxi  | Mosu | –                   | love  | –    | Kodo   | Bola |   |
| a  | mali  | matu | Padê                | manse | kasô | Avo    | Faka |   |

Na tentativa de esclarecer a origem do processo, o autor observa alguns dados diacrônicos e alega: em um momento inicial da formação do protocrioulo do Golfo da Guiné, a “aglutinação era fortemente associada à altura da vogal”<sup>8</sup> (cf. HAGEMEIJER, 2009 c: 38). Constatava-se a seguinte relação, não absoluta, entre as vogais das palavras originais, advindas do Português Europeu, e os segmentos aglutinados na protolíngua:

**Quadro 3** Aglutinação de vogais em palavras dissilábicas do Protocrioulo do Golfo da Guiné (cf. HAGEMEIJER, 2009c)

| VOGAIS TÔNICAS NAS PALAVRAS ORIGINAIS (PE) | VOGAIS INICIAIS AGLUTINADAS NAS PALAVRAS FINAIS (PCGG) | EXEMPLOS   |   |
|--|--|--|---|
|  |  | PE   | PCGG  |
| Baixa e médias-baixas<br>[a e ɔ]           | Baixa e média-baixa posterior<br>[a ɔ]                 | “Pá” – [ˈpa]<br>“fé” – [ˈfe]<br>“mar” – [ˈmar]<br>“pé” – [ˈpe]<br>“nó” – [ˈno] | [a]pa<br>[a]fe<br>[ɔ]mali<br>[ɔ]pe<br>[ɔ]no |
| [e i]                                      | [o]  | “céu” – [ˈsew]<br>“rio” – [ˈriw]   | [o]sê;<br>[o]lhô                            |
| Alta<br>[u]                                | Alta<br>[u]  | “nu” – [ˈnu]   | [u]nu                                       |

Além da restrição fonológica, motivadora da correspondência entre a altura da vogal tônica da palavra-base e a qualidade da vogal inicial aglutinada na protolíngua, os artigos definidos de gênero feminino (“a”) e masculino (“o”) parecem restringir essa aglutinação às vogais baixa [a], média baixa, média alta e alta posteriores [ɔ ou u]. A comparação entre os sistemas decorrentes da expansão da língua-mãe no tempo e no espaço indica, entretanto, comportamentos singulares. Na tentativa de explicá-los, Hagemeijer (2009 c: 43) adverte:

Inicialmente, o protocrioulo do Golfo da Guiné era essencialmente o resultado da aquisição do Português por escravos edóides. Durante este período correspondente, em grande medida, à sociedade de habitação, houve indiscutivelmente facilidade de acesso à língua alvo (TL) e, portanto, o decalque

<sup>8</sup> Tradução de “agglutination is strongly associated to vowel height”.

de gênero do Português foi também capaz de satisfazer a restrição edóide de iniciar os nomes por vogal.<sup>9</sup>

Atualmente, os crioulos falados na Ilha de São Tomé (Forro e Angolar) possuem um número considerável de palavras introduzidas por consoantes, indicando uma maior influência da língua de superstrato e a consequente baixa produtividade da aglutinação de vogais iniciais. No Forro, o fenômeno limita-se aos itens aglutinados naquele estágio diacrônico da protolíngua.

Em suma, a assimilação e a harmonia vocálicas (de altura e/ou ATR) são fenômenos comuns ao PE, PB e a todos crioulos do Golfo da Guiné (cf. HAGEMIJER, 2009: 37) que coexistem na Ilha com o Português, sendo provável que também se manifestem na amostra investigada.

### 3. ANÁLISE DOS DADOS

A amostra geral que serve de base à presente investigação integra 11.179 ocorrências, selecionadas dos inquéritos que constituem o Projeto *Varietades Africanas do Português (VAPOR)*, do Centro de Linguística da Universidade de Lisboa. Trata-se de 6.643 dados de /e/ e 4.536 de /o/, extraídos de Diálogos entre Informante e Documentador (DID), realizados com dezessete indivíduos, residentes na Ilha de São Tomé e distribuídos por sexo, três faixas etárias (25-35 anos; 36-55 anos; 56 anos em diante) e três níveis de escolaridade (Fundamental, Médio e Superior). Com base nas auto declarações dos entrevistados, avaliou-se, também, a frequência de uso do Forro na aplicabilidade do alçamento. Finalmente, postularam-se onze variáveis estruturais, a saber: qualidade da vogal da sílaba seguinte, nasalidade da vogal alvo, características articulatórias (pontos e modos de articulação) das consoantes dos ataques precedente e subsequente à pretônica, tipo de sílaba, distância entre a vogal alvo e outra alta presente no vocábulo, localização da pretônica na palavra e classe gramatical.

A princípio, relacionaram-se cinco variantes das médias pretônicas (alta, média-alta, média-baixa, ditongada e zero) às informações estruturais e sociais supracitadas. A baixa produtividade dos processos de abertura, de ditongação e de apagamento, nas duas amostras (cf. Tabela 2), implicou o descarte de tais

<sup>9</sup> Tradução de “early proto-GGC is essentially the result of Edoid slaves acquiring Portuguese. Since this period corresponds by and large to the *société d’habitation*, there was arguably better access to the Target Language (TL) and therefore phonetic calquing upon Portuguese gender was additionally able to satisfy the Edoid constraint that nouns are vowel-initial”.

ocorrências das análises variacionistas e a sua centralização no alçamento e na manutenção do timbre médio.

**Tabela 2** Variantes das vogais médias pretônicas /e/ e /o/ no PST

| VARIANTES            | /e/        |            | /o/        |             |
|----------------------|------------|------------|------------|-------------|
|                      | OCO.       | %          | OCO.       | %           |
| [i u]                | 3.371      | 50.8       | 2.193      | 48.3        |
| [e o]                | 2.544      | 38.3       | 1.736      | 38.3        |
| [e ɔ]                | <b>68</b>  | <b>1</b>   | <b>27</b>  | <b>0.6</b>  |
| [ej oj]              | <b>266</b> | <b>4</b>   | <b>558</b> | <b>12.3</b> |
| [ ]                  | <b>394</b> | <b>5.9</b> | <b>22</b>  | <b>0.5</b>  |
| Total de ocorrências | 6.643      | 59.4       | 4.536      | 40.6        |

Eliminaram-se, também, os casos pouco produtivos e semicategóricos/categóricos e modificou-se a configuração inicial de determinados grupos de fatores. Após as alterações, a elevação foi estabelecida como valor de aplicação e os dados expostos a rodadas no *GoldVarb X*.

### 3.1 A pretônica /e/

Reiterando o exposto acima, 59.4% da amostra geral, ou seja, 6.643 dados referem-se à vogal /e/. Concentrar a investigação nas pronúncias [i] e [e] reduziu o número inicial a 5.915 casos, distribuídos por variante conforme a Tabela 3.

**Tabela 3** Distribuição das ocorrências de alçamento e manutenção da vogal /e/

| VARIANTES | OCO.  | %  |
|-----------|-------|----|
| [i]       | 3.371 | 57 |
| [e]       | 2.544 | 43 |

Postularam-se as hipóteses de maior incidência da pronúncia alta quando a vogal está no início de palavra: i) em sílabas desprovidas de ataque e de coda; ii) naquelas, cuja posição de ataque está disponível e a coda é preenchida por /S/ ou /N/; e iii) na sequência /deS/ (prefixal ou não). Além desses ambientes, cogitou-se a maior probabilidade de elevação nas vogais inseridas em contextos de hiato; e

nos itens “depois”, “pessoa(s)” e “pessoal”. Como se verifica na Tabela 4, só no contexto de hiato, a hipótese inicial não se confirmou:

**Tabela 4** Distribuição das ocorrências de [e] e de [i], no PST, em alguns contextos aparentemente mais suscetíveis à elevação

|                                |                           | [i]  |             | [e]  |             |
|--------------------------------|---------------------------|------|-------------|------|-------------|
|                                |                           | OCO. | %           | OCO. | %           |
| Em início absoluto de vocábulo | /eS/                      | 922  | <b>99.5</b> | 5    | 0.5         |
|                                | /eN/                      | 278  | <b>70.4</b> | 117  | 29.6        |
|                                | /e/                       | 216  | <b>71</b>   | 89   | 29.2        |
|                                | /deS/ prefixal ou não     | 137  | <b>98.6</b> | 2    | 1.4         |
| Outros casos                   | Hiato                     | 45   | 34.6        | 85   | <b>65.4</b> |
|                                | “D/e/pois”                | 177  | <b>60</b>   | 118  | 40          |
|                                | “p/e/ssoa(s), p/e/ssoaal” | 344  | <b>98.6</b> | 5    | 1.4         |

Além de haver menos registros de advérbios, numerais e conjunções com a vogal /e/, neles o alçamento se destacou. Uma análise qualitativa dos resultados comprova, contudo, a interferência da repetição de certos itens. Além dessas ocorrências, desprezaram-se, por comportamento semicategórico/categórico em favor da pronúncia [i], os itens isolados “pessoa(s)”, “pessoal” e “depois”. A amostra final avaliou, então, 3.334 ocorrências de vogal média não recuada, distribuídas entre as variantes [i] e [e] do seguinte modo:

**Tabela 5** Redistribuição das ocorrências de alçamento e manutenção da vogal /e/

| PRETÔNICA ANTERIOR | OCO.  | %    |
|--------------------|-------|------|
| [i]                | 1.216 | 36.5 |
| [e]                | 2.118 | 63.5 |

A rodada eleita como mais coerente estabeleceu algumas variáveis como influentes na elevação de /e/. Cada uma delas será descrita nas subseções adiante, comparando-se, sempre que possível, os resultados aqui obtidos àqueles divulgados por Rocha (2013), relativos a dados contemporâneos do PB. Por ora, o Quadro 4 destaca esses condicionamentos em uma escala decrescente de relevância, bem como a significância e o *input* da regra na variedade urbana do Português de São Tomé.

**Quadro 4** Rodada selecionada para a elevação de /e/

| <b>ELEVAÇÃO DA PRETÔNICA /e/</b>                                      |                            |
|---|----------------------------|
| <i>Qualidade da vogal da sílaba seguinte</i>                          |                            |
| <i>Modo de articulação da consoante precedente</i>                    |                            |
| <i>Modo de articulação da consoante do ataque da sílaba seguinte</i>  |                            |
| <i>Ponto de articulação da consoante do ataque da sílaba seguinte</i> |                            |
| <i>Distância entre a vogal alvo e outra alta presente no vocábulo</i> |                            |
| <i>Classe gramatical do vocábulo</i>                                  |                            |
| <i>Grau de escolaridade dos Informantes</i>                           |                            |
| <i>Natureza da atonicidade</i>  |                            |
| <i>Faixa etária dos informantes</i>                                   |                            |
| <i>Sexo dos informantes</i>   |                            |
| <i>Frequência de uso do Forro</i>                                     |                            |
| <b>Input: .42</b>   | <b>Significance: 0.010</b> |

a) *Qualidade da vogal da sílaba seguinte*

O condicionamento mais determinante para o alteamento de /e/ na amostra investigada é a qualidade da vogal da sílaba contígua à pretônica, cujos fatores cogitados e seus respectivos índices distribuem-se conforme a Tabela 6:

**Tabela 6** Atuação da variável *Tipo de vogal contígua* para o alçamento de /e/, no PST

| <b>VARIÁVEL TIPO DE VOGAL CONTÍGUA</b>                 |  | <b>[i]</b>  |           |              |
|--|--|-------------|-----------|--------------|
|  |  | <b>OCO.</b> | <b>%</b>  | <b>P. R.</b> |
| <b>Vogal tônica/átona<br/>(oral/nasal/ nasalizada)</b> | <b>Altas (homorgânica e não homorgânica)</b> | 550/1.088   | <b>51</b> | <b>.55</b>   |
|  | <b>Médias-altas</b>                          | 457/1.290   | <b>35</b> | <b>.54</b>   |
|  | <b>Baixas (médias-baixas e baixa)</b>        | 209/956     | 22        | .39          |

A princípio, cogitou-se a influência das diferentes alturas da vogal contígua e, ainda, da homorganicidade e da não homorganicidade das altas, na manifestação do alçamento da pretônica /e/. No entanto, a constatação de comportamentos semelhantes permitiu redistribuí-las em três graus de abertura: altas (homorgânicas e não homorgânicas), médias-altas e baixas (médias-baixas e baixa). Embora com pesos relativos muito próximos da neutralidade (P. R.: .50), a pronúncia [i] é mais provável, se na sílaba contígua há uma vogal alta (homorgânica ou não) (P. R.: .55) ou vogais médias altas (P. R.: .54). Por outro lado, a contiguidade com segmentos baixos favorece a manutenção de timbre em /e/ (P. R.: .39).

Em Rocha (2013), a qualidade da vogal contígua também foi selecionada como o fator mais influente na elevação da média anterior. Como ilustra a Tabela 7, a autora considera aspectos mais específicos nesta variável. As peculiaridades entre as investigações não impedem, contudo, que se estabeleçam comparações. Nas duas realidades, as altas se destacam, sendo comprovada a elevação da média pretônica por harmonização com o traço alto da vogal contígua em Nova Iguaçu e em São Tomé.

**Tabela 7** Atuação da variável *Tipo de vogal contígua* para o alçamento de /e/, em Nova Iguaçu (cf. ROCHA, 2013: 111)

| VARIÁVEL TIPO DE VOGAL CONTÍGUA         |                      | [i]       |    |       |
|---|----------------------|-----------|----|-------|
|   |                      | OCO.      | %  | P. R. |
| Vogal tônica<br>(oral/nasal/nasalizada) | Alta homorgânica     | 460/882   | 52 | .95   |
|   | Alta não homorgânica | 87/172    | 51 | .73   |
|   | Não altas            | 196/3.543 | 5  | .29   |
| Vogal átona<br>(oral/nasal/nasalizada)  | Alta homorgânica     | 104/550   | 19 | .81   |
|   | Alta não homorgânica | 29/184    | 16 | .74   |
|   | Não altas            | 70/882    | 8  | .31   |

*b) Modo de articulação da consoante precedente*

O segundo fator mais influente na elevação de /e/ diz respeito ao processo de redução vocálica e foi cogitado na intenção de avaliar a relação entre algumas consoantes e o processo de alçamento, isto é, de verificar a ocorrência de assimilação entre consoantes e vogais nas amostras.

**Tabela 8** Atuação da variável *Modo de articulação da consoante precedente* para o alçamento de /e/, no PST

| VARIÁVEL MODO DE ARTICULAÇÃO<br>(CONTEXTO PRECEDENTE) | [i]       |           |            |
|---|-----------|-----------|------------|
|   | OCO.      | %         | P. R.      |
| Oclusivas   | 335/1.001 | 33.5      | .49        |
| <b>Fricativas</b>                                     | 498/1.017 | <b>49</b> | <b>.66</b> |
| Líquidas  | 208/855   | 24        | .33        |
| Nasais  | 168/448   | 37.5      | .47        |

Como aponta a Tabela 8, na análise final, esta variável considerou apenas os dados de segmentos oclusivos, fricativos, nasais e líquidos (laterais e vibrantes), pois os casos de ataque vazio se tornaram escassos à proporção que se descartaram as ocorrências em início absoluto com coda vazia ou preenchida por /S/ ou /N/.

Todos os índices percentuais dos contextos preservados, confirmam a menor incidência da elevação. Já os pesos relativos indicam que, quando /e/ atua como núcleo de sílabas com ataque preenchido por consoantes fricativas, a pronúncia [i] é mais provável (P. R.: .66); e que a manutenção se destaca sucedendo líquidas (P. R.: .33). Por outro lado, segmentos oclusivos (P. R.: .49) e nasais (P. R.: .47) apresentaram índices muito próximos da neutralidade, mesmo mais favoráveis à manutenção.

Se comparados os *corpora* contemporâneos de São Tomé e de Nova Iguaçu, verifica-se que os dois estudos descartam os casos de ataque vazio, em virtude de a maioria das ocorrências se referir às sequências /e/, /eN/ e /eS/ em início de vocábulo. No PB, o primeiro contexto (/e/) seria mais suscetível à manutenção e a elevação mais provável nos outros dois (/eN/ e /eS/) (cf. BRANDÃO et al., 2012). Já na fala urbana de São Tomé, além das sequências /eN/ e /eS/, sílabas leves desprovidas de ataque tendem ao alçamento, sendo, portanto, a pronúncia [i] mais frequente nos três ambientes.

No que concerne aos demais modos de articulação precedentes, há peculiaridades entre as duas amostras. Todos os percentuais indicam o predomínio de [e]. No primeiro estudo, consoantes nasais (20%, P. R.: .53) e líquidas (14%, P. R.: .53) se mostraram, contudo, mais suscetíveis à elevação, em contraste com as oclusivas (9%, P. R.: .49) e fricativas (12%, P. R.: .48). Já neste, segmentos fricativos (49%, P. R.: .66) estão mais sujeitos ao alçamento do que líquidos (24%, P. R.: .33), oclusivos (33.5%, P. R.: .49) e nasais (37.5%, P. R.: .47).

### c) *Modo de articulação da consoante do ataque da sílaba seguinte*

No que se refere ao modo de articulação da consoante que ocupa o ataque da sílaba subsequente à vogal alvo, os dados de [i] se distribuem conforme a Tabela 9:

**Tabela 9** Atuação da variável *Modo de articulação da consoante do ataque subsequente* para o alçamento de /e/, no PST

| VARIÁVEL MODO DE ARTICULAÇÃO<br>(ATAQUE SUBSEQUENTE) | [i]       |    |            |
|--|-----------|----|------------|
|  | OCO.      | %  | P. R.      |
| Oclusivas  | 426/1.335 | 32 | .42        |
| <b>Nasais</b>  | 120/251   | 48 | <b>.56</b> |
| <b>Fricativas</b>                                    | 489/1.145 | 43 | <b>.58</b> |
| <b>Laterais</b>                                      | 95/194    | 49 | <b>.54</b> |

Todos os percentuais acima remetem à maior incidência da manutenção, porém os pesos relativos indicam que a presença de fricativas (P. R.: .58), nasais (P. R.: .56) e laterais (P. R.: .54), respectivamente, no início da sílaba que sucede a pretônica contribuem para sua elevação, ao passo que a pronúncia [e] é mais provável quando a mesma posição é preenchida por oclusivas (P. R.: .42).

A Tabela 10 compara os resultados de Rocha (2013) aos aqui expostos para a mesma variável:

**Tabela 10** Atuação da variável *Modo de articulação da consoante do ataque subsequente* para o alçamento de /e/, em Nova Iguaçu (cf. ROCHA, 2013: 113) e no PST

| VARIÁVEL MODO DE ARTICULAÇÃO<br>(ATAQUE SUBSEQUENTE) | [i]            |    |       |           |    |       |
|--|----------------|----|-------|-----------|----|-------|
|  | NOVA IGUAÇU/RJ |    |       | SÃO TOMÉ  |    |       |
|  | OCO.           | %  | P. R. | OCO.      | %  | P. R. |
| Oclusivas  | 242/1.795      | 13 | .73   | 426/1.335 | 32 | .42   |
| Fricativas   | 277/2.475      | 11 | .38   | 489/1.145 | 43 | .58   |
| Líquidas   | 97/1.165       | 8  | .27   | –         | –  | –     |
| Nasais   | 221/551        | 40 | .71   | 120/251   | 48 | .56   |
| Laterais   | –              | –  | –     | 95/194    | 49 | .54   |

Há certa correspondência entre as duas variedades, pois, nas duas amostras, consoantes nasais subsequentes à pretônica favorecem a realização alta (40%, P. R.: .71 em Nova Iguaçu; e 48%, P. R.: .56 em São Tomé). Ao mesmo tempo, se, naquele estudo, oclusivas apresentam percentual de [i] superior (13%) ao de fricativas (11%); neste, o contrário se verifica, tendo os segmentos fricativos índices maiores (43%) do que oclusivos (32%).

d) *Ponto de articulação da consoante do ataque da sílaba seguinte*

O quarto contexto de provável elevação da pretônica anterior é o ponto de articulação da consoante do ataque subsequente. Esta variável foi modificada, reorganizando-se de acordo com as variantes expressas na Tabela 11 e todos os percentuais comprovam a maior aplicabilidade da regra de manutenção. Quanto aos pesos relativos, ambos se aproximam da neutralidade. No entanto, precedendo consoantes [+coronais], a tendência à pronúncia [i] é maior (P. R.: .52). Vale destacar, ainda, a discrepância entre os dados de cada um dos contextos em evidência, já que segmentos [+coronais] (2.092) são mais numerosos na amostra, se comparados aos [-coronais] (1.141).

**Tabela 11** Atuação da variável *Ponto de articulação da consoante do ataque subsequente* para o alçamento de /e/, no PST

| VARIÁVEL PONTO DE ARTICULAÇÃO<br>(ATAQUE SUBSEQUENTE) | [i]       |      |       |
|---|-----------|------|-------|
|   | OCO.      | %    | P. R. |
| [+ Coronais] (alveolares e alveopalatais)             | 743/2.092 | 35.5 | .52   |
| [-coronais] (labiais e velares)                       | 383/1.141 | 33.5 | .47   |

Rocha (2013) também assinala esta variável como influente nos processos de manutenção e de alçamento da vogal média anterior em Nova Iguaçu e descreve resultados semelhantes, mesmo postulando fatores distintos. Como ilustra a Tabela 12, naquela análise, as consoantes do ataque subsequente se distribuem em [+coronais, +anteriores], [+coronais, -anteriores] e [-coronais]. Ressaltam-se, contudo, percentuais de elevação i) superiores no segundo grupo (18%); e ii) idênticos, quando a vogal /e/ precede segmentos [+coronais, +anteriores] e [-coronais] (15%).

**Tabela 12** Atuação da variável *Ponto de articulação da consoante do ataque subsequente* para o alçamento de /e/, em Nova Iguaçu (cf. ROCHA, 2013: 119)

| VARIÁVEL PONTO DE ARTICULAÇÃO<br>(ATAQUE SUBSEQUENTE) | [i]       |    |       |
|---|-----------|----|-------|
|   | OCO.      | %  | P. R. |
| [+ Coronais, + anteriores] (alveolares)               | 548/3.703 | 15 | .56   |
| [+ Coronais, -anteriores] (Alveopalatais e palatais)  | 70/384    | 18 | .56   |
| [-Coronais] (labiais e velares)                       | 328/2.125 | 15 | .37   |

e) *Distância entre a vogal alvo e outra alta presente no vocábulo*

A distância entre a vogal alvo e outra alta presente na palavra foi o quinto aspecto relevante para o alçamento de /e/ na variedade urbana do PST. Compõem o grupo, o número de sílabas entre o segmento alvo e outro alto acentuado ou não.

Trata-se de uma variável que dialoga com a qualidade da vogal da sílaba seguinte, ressaltando o acento como fator relevante: a pronúncia [i] é mais recorrente em presença de um segmento alto tônico e contíguo (P. R.: .64). Em contrapartida, i) quando há altas átonas, não contíguas (P. R.: .22) e contíguas (P. R.: .38), a manutenção é mais frequente; e ii) tônicas não contíguas aproximam-se da neutralidade (P. R.: .41), mesmo tendendo à manutenção.

**Tabela 13** Atuação da variável *Distância de outra vogal alta na palavra* para o alçamento de /e/, no PST

| VARIÁVEL <i>DISTÂNCIA DE OUTRA VOGAL ALTA</i> |              | [i]     |      |            |
|---|--------------|---------|------|------------|
|   |              | OCO.    | %    | P. R.      |
| Tônica  | Contígua     | 400/675 | 59   | <b>.64</b> |
|   | Não contígua | 55/161  | 34   | .41        |
| Átona   | Contígua     | 150/418 | 36   | .38        |
|   | Não contígua | 26/98   | 26.5 | .22        |

f) *Classe gramatical dos vocábulos*

O sexto fator selecionado como favorável à elevação de /e/ foi a classe gramatical dos vocábulos de cuja estrutura a pretônica faz parte. A variável foi submetida a uma série de alterações durante a investigação, sendo preservados, na abordagem variacionista, apenas os fatores apontados na Tabela 14:

**Tabela 14** Atuação da variável *Classe de vocábulos* para o alçamento de /e/, no PST

| VARIÁVEL <i>CLASSES DE VOCÁBULOS</i> | [i]       |    |            |
|--------------------------------------|-----------|----|------------|
|                                      | OCO.      | %  | P. R.      |
| Substantivos                         | 410/1.276 | 32 | .48        |
| Adjetivos                            | 149/494   | 30 | .44        |
| <b>Verbos: formas finitas</b>        | 180/564   | 32 | <b>.51</b> |
| <b>Verbos: formas não finitas</b>    | 317/740   | 43 | <b>.56</b> |

Como se verifica, todos os percentuais sugerem a maior incidência da pronúncia [e] na amostra com pesos relativos mais próximos da neutralidade. A elevação é, todavia, mais provável nos verbos, sobretudo em suas formas não finitas (P. R.: .56). Em contrapartida, adjetivos (P. R.: .44) e substantivos (P. R.: .48) estão mais sujeitos à retenção do timbre médio na série anterior.

g) *Grau de escolaridade dos informantes*

As quatro variáveis extralinguísticas consideradas nesta investigação, a saber: grau de escolaridade, faixa etária, sexo dos informantes e frequência de uso do Forro, foram selecionadas pela análise estatística, correspondendo, nessa

ordem, às sétima, nona, décima e décima primeira variáveis influentes no alçamento de /e/. O nível de escolaridade foi postulado na tentativa de avaliar o prestígio das pronúncias média-alta e alta no dialeto em pauta. Como é possível depreender da Tabela 15, todos os pesos relativos se aproximam da neutralidade (P. R.: .50), mas em informantes com nível superior, a pronúncia [i] é mais provável (P. R.: .56), aqueles com nível médio tendem à manutenção (P. R.: .44) e os de nível Fundamental são neutros (P. R.: .50).

**Tabela 15** Atuação da variável *Escolaridade dos informantes* para o alçamento de /e/, no PST

| VARIÁVEL ESCOLARIDADE DOS INFORMANTES             | [i]       |      |            |
|---|-----------|------|------------|
|   | OCO.      | %    | P. R.      |
| Fundamental – 5 a 8 anos de escolaridade          | 252/818   | 31   | .50        |
| Médio – 9 a 11 anos de escolaridade               | 405/1.212 | 33.5 | .44        |
| <b>Superior – mais de 11 anos de escolaridade</b> | 559/1.304 | 43   | <b>.56</b> |

*b) Natureza da atonicidade*

A oitava variável selecionada como determinante para a aplicabilidade da regra de elevação em /e/ foi a natureza da atonicidade da vogal alvo. A variável foi explorada, a fim de observar se, na amostra investigada, a pronúncia [i] é mais frequente quando integra o núcleo de sílabas sempre átonas nas diferentes formas de um mesmo paradigma (cf. BISOL, 1981: 82). Os índices percentuais descritos na Tabela 16 revelam a baixa produtividade do alçamento nos dois tipos de vogal átona considerados (36% para eventuais e 37% para permanentes). No entanto, corroborando a perspectiva de Bisol (1981) sobre o PB, na fala urbana do Português de São Tomé, a pronúncia [i] é mais provável em sílabas sempre átonas (P. R.: .59).

**Tabela 16** Atuação da variável *Natureza da atonicidade da vogal alvo* para o alçamento de /e/, no PST

| VARIÁVEL NATUREZA DA ATONICIDADE DE /E/ | [i]       |    |            |
|---|-----------|----|------------|
|   | OCO.      | %  | P. R.      |
| Átona eventual                          | 870/2.400 | 36 | .46        |
| <b>Átona permanente</b>                 | 346/934   | 37 | <b>.59</b> |

i) *Faixa etária dos informantes*

A faixa etária dos informantes foi também selecionada como motivação da pronúncia [i] na amostra. À semelhança de outras pesquisas de cunho variacionista, foi postulada para averiguar se há indícios de mudança ou se a flutuação fonética entre manutenção e alçamento nesta realidade linguística permanece estável. Sobre sua interferência na aplicabilidade da regra de elevação em Nova Iguaçu, Rocha (2013: 121) adverte:

os percentuais e ocorrências (...) evidenciam a maior frequência da retenção do timbre médio (...), independentemente da geração, já que, em todas as faixas etárias ressaltadas, essa pronúncia supera os 80%. Os pesos relativos indicam, contudo, que indivíduos mais jovens são mais resistentes à mudança sonora de /e/ (.46), ao passo que os de meia idade e idosos igualmente se mostraram mais suscetíveis à sua aplicação (.52).

Na Tabela 17, apontam-se os índices registrados para a mesma variável na fala urbana do Português de São Tomé. Como se verifica, os percentuais e ocorrências descritos comprovam a maior incidência de [e] e, indicando um cenário de variação estável, todos os pesos relativos se aproximam da neutralidade, mesmo revelando a maior resistência de indivíduos mais velhos (P. R.: .46) e mais jovens (P. R.: .49) à elevação, enquanto que na faixa intermediária o processo seria mais provável (P. R.: .54).

**Tabela 17** Atuação da variável *Faixa etária dos informantes* para o alçamento de /e/, no PST

| VARIÁVEL FAIXA ETÁRIA | [i]       |      |            |
|-----------------------|-----------|------|------------|
|                       | OCO.      | %    | P. R.      |
| 25 – 35 anos          | 502/1.359 | 37   | .49        |
| <b>36 – 55 anos</b>   | 446/1.107 | 40.5 | <b>.54</b> |
| 56 anos em diante     | 268/868   | 31   | .46        |

j) *Sexo dos Informantes*

O sexo dos informantes foi a penúltima variável selecionada como relevante para o alçamento de /e/ na amostra. A variável foi inserida entre as demais informações extralinguísticas dos entrevistados, com o fito de constatar se a vogal pretônica varia na fala de homens e de mulheres no dialeto em pauta. Estudos sociolinguísticos costumam incluí-la no perfil dos informantes, salientando que, de acordo com a avaliação social da variante, a mudança sonora é mais provável em homens do que em mulheres. Ressaltam, entretanto, a maior adequação da

fala feminina à norma de prestígio. Na variedade urbana do Português de São Tomé, a variante alta se mostrou pouco produtiva em ambos os sexos, com percentuais inferiores a 40%, sendo mais representativa em homens.

**Tabela 18** Atuação da variável *Sexo dos informantes* para o alçamento de /e/, no PST

| VARIÁVEL SEXO DOS INFORMANTES | [i]       |      |       |
|-------------------------------|-----------|------|-------|
|                               | OCO.      | %    | P. R. |
| Masculino                     | 805/2.144 | 37.5 | .53   |
| Feminino                      | 411/1.190 | 34.5 | .44   |

Em consonância com os índices percentuais e de ocorrências, os pesos relativos expressos na Tabela 18 indicam, assim, que a variante [i] é mais frequente no discurso masculino (P. R.: .53) do que no feminino (P. R.: .44).

*k) Frequência de uso do Forro*

A última variável selecionada como influente na elevação de /e/ no *corpus* em destaque foi a frequência de uso do Forro. Reconhece-se o multilinguismo como inerente à realidade santomense e a presente variável foi inserida nas análises na tentativa de investigar a interferência do contato entre o sistema autóctone mais falado na região e o Português. Tendo em vista os relatos dos entrevistados, consideram-se, três diferentes graus de emprego do crioulo:

**Tabela 19** Atuação da variável *Frequência de uso do Forro* para o alçamento de /e/, no PST

| VARIÁVEL FREQUÊNCIA DE USO DO FORRO | [i]       |      |       |
|-------------------------------------|-----------|------|-------|
|                                     | OCO.      | %    | P. R. |
| Pouca ou nenhuma                    | 590/1.536 | 38.5 | .50   |
| Média                               | 572/1.559 | 36.5 | .53   |
| Alta                                | 54/239    | 22.5 | .34   |

A comparação entre os valores expressos na Tabela 19 revela que a maioria das ocorrências concentra-se entre indivíduos que declararam pouca/nenhuma e média frequência do Forro em seu cotidiano. Apenas um dos dezessete entrevistados declarou alta frequência, o que justifica a menor produtividade do alçamento nesse quesito (P. R.: .34). Vale ressaltar, por fim, que, embora as duas outras possibilidades se aproximem da neutralidade, indivíduos cujo contato

com o Forro é mediano estão mais propensos à elevação da pretônica anterior (P. R.: .53).

### 3.2 A pretônica /o/

A pretônica /o/ está representada em cerca de 40% (4.536 dados) da amostra geral. A Tabela 20 retoma o exposto no início deste capítulo, enfatizando as cinco variantes pelas quais se distribuem os dados da série em pauta, representativas, respectivamente, dos processos de alçamento, manutenção, abaixamento, ditongação e apagamento:

**Tabela 20** Distribuição das ocorrências pelas variantes de /o/, no PST

| VARIANTES | OCO.  | %    |
|-----------|-------|------|
| [u]       | 2.193 | 48.3 |
| [o]       | 1.736 | 38.3 |
| [ɔ]       | 27    | 0.6  |
| [ow]      | 558   | 12.3 |
| [ ]       | 22    | 0.5  |

As variantes baixa, ditongada e zero foram pouco produtivas. Por isso, optou-se por centralizar a análise final nos dois primeiros processos. A restrição reduziu o *corpus* a 3.929 dados, cerca de 56% de [o] e 44% de [u].

**Tabela 21** Distribuição das ocorrências pelas variantes média fechada e alta de /o/, no PST

| VARIANTE | OCO.  | %    |
|----------|-------|------|
| [u]      | 2.193 | 55.8 |
| [o]      | 1.736 | 44.2 |

A observação mais específica da variável *Tipo de sílaba* demonstrou que, descartados os dados de abertura, de ditongação e de apagamento, os contextos de travamento por /l/, /N/ e /S/ foram pouco produtivos. As ocorrências de sílabas sem travamento em contexto de hiato são, de igual modo, reduzidas. Nelas, a elevação é, todavia, recorrente, assim como nos casos de coda preenchida por /R/.

**Tabela 22** Distribuição das ocorrências de manutenção e de alçamento de /o/, por tipo de sílaba, no PST

| VARIÁVEL TIPO DE SÍLABA             | [u]   |      | [o]   |      |
|-------------------------------------|-------|------|-------|------|
|                                     | OCO.  | %    | OCO.  | %    |
| Travada por /S/                     | 77    | 64.2 | 43    | 35.8 |
| Travada por /N/                     | 41    | 23.3 | 135   | 76.7 |
| Travada por /R/                     | 794   | 72.8 | 297   | 27.2 |
| Travada por //                      | 5     | 15.2 | 28    | 84.8 |
| Sem travamento                      | 1.177 | 49.4 | 1.207 | 50.6 |
| Sem travamento em contexto de hiato | 98    | 82.4 | 21    | 17.6 |

Como ilustra a Tabela 23, a alta frequência dos itens lexicais “porque” (517/793) e “portanto” (123/793) parece interferir nos resultados do último grupo. Em tais vocábulos, em “bocado(s), bocadinho” e em “Tomé, tomense(s)”, a pronúncia [u] é frequente, ao passo que em “você”, [o] se destaca. Os poucos dados de “comigo/consigo” variam entre as duas pronúncias.

**Tabela 23** Distribuição das ocorrências de elevação e manutenção de timbre, de acordo com o item analisado no PST

| ITEM LEXICAL           | [u]  |            | [o]  |           |
|------------------------|------|------------|------|-----------|
|                        | OCO. | %          | OCO. | %         |
| “Porque”               | 511  | <b>99</b>  | 6    | 1         |
| “Portanto”             | 122  | <b>99</b>  | 1    | 1         |
| “Bocado(s), bocadinho” | 80   | <b>100</b> | 0    | 0         |
| “Tomé, tomense(s)”     | 199  | <b>93</b>  | 14   | 7         |
| “Comigo, consigo”      | 7    | <b>44</b>  | 9    | <b>56</b> |
| “Você”                 | 4    | 12         | 30   | <b>88</b> |

As palavras supracitadas foram desconsideradas na abordagem variacionista, para que o condicionamento lexical não interferisse nos resultados. As eliminações reduziram os dados de conectores, numerais e advérbios (em geral e em -mente) a poucos itens lexicais e ocorrências de manutenção e de alçamen-

to. Optou-se, portanto, por desconsiderá-los, limitando a última amostra a 2.693 ocorrências de vogal média recuada, distribuídas conforme a Tabela 24:

**Tabela 24** Distribuição das ocorrências de vogal /o/ por variante considerada no PST

| PRETÔNICA POSTERIOR | OCO.  | %    |
|---------------------|-------|------|
| [u]                 | 1.144 | 42.5 |
| [o]                 | 1.549 | 57.5 |

No Quadro 5, listam-se as motivações do alçamento da pretônica /o/, com base na rodada selecionada para a descrição dos resultados. Em consonância com o efetuado para o tratamento de /e/, adiante comenta-se, por ordem de relevância, cada uma das variáveis apontadas.

**Quadro 5** Rodada selecionada para /o/

| PRETÔNICA /o/   |                             |
|---|-----------------------------|
| <i>Modo de articulação da consoante precedente</i>                    |                             |
| <i>Distância entre a vogal alvo e outra alta presente no vocábulo</i> |                             |
| <i>Grau de escolaridade dos informantes</i>                           |                             |
| <i>Qualidade da vogal da sílaba seguinte</i>                          |                             |
| <i>Modo de articulação da consoante do ataque da sílaba seguinte</i>  |                             |
| <i>Nasalidade da vogal alvo</i>                                       |                             |
| <i>Classe gramatical do vocábulo</i>                                  |                             |
| <i>Frequência de uso do Forro</i>                                     |                             |
| <i>Sexo dos informantes</i>   |                             |
| <i>Ponto de articulação da consoante precedente</i>                   |                             |
| <i>Faixa etária dos informantes</i>                                   |                             |
| <b>Input: .33</b>   | <b>Significância: 0.017</b> |

a) *Modo de articulação da consoante precedente*

O modo de articulação da consoante precedente foi o primeiro fator selecionado como relevante para a elevação de /o/ na variedade urbana do Português de São Tomé.

Trata-se de uma das variáveis que remete ao processo de redução vocálica e que visa a estabelecer relações de assimilação entre traços vocálicos e consonantais. Os casos de vogal precedida por consoantes laterais e vibrantes foram poucos e apresentam comportamentos muito semelhantes no que tange à distribuição

das variantes média e alta, justificando seu agrupamento em um único fator: as consoantes líquidas. À exceção das oclusivas que correspondem ao maior número de ocorrências (1.467) e tendem à elevação (55% e P. R.: .61), todos os índices apontados na Tabela 25 comprovam a manutenção de /o/ como regra no dialeto, independentemente do contexto precedente, com pesos .31, .33 e .39, respectivamente, em presença de segmentos fricativos, nasais e líquidos.

**Tabela 25** Atuação da variável *Modo de articulação da consoante precedente* para o alçamento de /o/, no PST

| VARIÁVEL MODO DE ARTICULAÇÃO<br>(CONTEXTO PRECEDENTE) | [u]       |    |       |
|---|-----------|----|-------|
|   | OCO.      | %  | P. R. |
| Oclusivas   | 806/1.467 | 55 | .61   |
| Fricativas  | 88/329    | 27 | .31   |
| Líquidas (laterais e vibrantes)                       | 154/462   | 33 | .39   |
| Nasais  | 81/276    | 29 | .33   |

Em Rocha (2013), a variável em questão foi também selecionada como influente na elevação de /o/. Apesar de os percentuais comprovarem a manutenção como regra no dialeto (cf. Tabela 26), à semelhança da amostra em evidência, “na fala de Nova Iguaçu, consoantes oclusivas (28.3%) (...) foram menos resistentes” (p. 125) ao alteamento. A comparação entre os índices expressos adiante confirma a importância de tais sons para a incidência da variante [u], tanto no Português do Brasil quanto no Português de São Tomé. Em contrapartida, sucedendo líquidas e nasais, a vogal em questão parece ser mais resistente à elevação, nas duas variedades.

**Tabela 26** Atuação da variável *Modo de articulação da consoante precedente* para o alçamento de /o/, em Nova Iguaçu (cf. ROCHA, 2013: 125) e no PST

| VARIÁVEL MODO DE ARTICULAÇÃO<br>(CONTEXTO PRECEDENTE) | [u]            |    |       |           |    |       |
|---|----------------|----|-------|-----------|----|-------|
|   | NOVA IGUAÇU/RJ |    |       | SÃO TOMÉ  |    |       |
|   | OCO.           | %  | P. R. | OCO.      | %  | P. R. |
| Oclusivas   | 724/2.559      | 28 | .64   | 806/1.467 | 55 | .61   |
| Fricativas  | 87/553         | 16 | .57   | 88/329    | 27 | .31   |
| Líquidas (laterais e vibrantes)                       | 22/740         | 3  | .13   | 154/462   | 33 | .39   |
| Nasais  | 41/396         | 10 | .38   | 81/276    | 29 | .33   |

b) *Distância entre a vogal alvo e outra alta presente no vocábulo*

O segundo fator influente para o alçamento de /o/ na amostra em pauta foi a distância entre a pretônica alvo e uma vogal alta tônica ou átona presente no vocábulo. Trata-se de uma variável que dialoga com a qualidade da vogal contígua, na avaliação da atuação do processo de harmonia vocálica, demonstrando que, quando na sílaba subsequente à média posterior há um segmento altoônico, a pronúncia [u] é mais provável (P. R.: .72).

**Tabela 27** Atuação da variável *Distância entre a pretônica e outra vogal alta na palavra* para o alçamento de /o/, no PST

| VARIÁVEL DISTÂNCIA DE OUTRA VOGAL ALTA |              | [u]     |      |       |
|--|--------------|---------|------|-------|
|  |              | OCO.    | %    | P. R. |
| Tônica                                 | Contígua     | 252/369 | 68   | .72   |
|  | Não contígua | 52/170  | 30.5 | .50   |
| Átona                                  | Contígua     | 278/764 | 36.5 | .40   |
|  | Não contígua | 28/109  | 26   | .43   |

Embora os pesos relativos das demais variantes dessa variável se aproximem da neutralidade, dividir a palavra com uma vogal alta átona contígua (P. R.: .40) ou não contígua (P. R.: .43) suscita a retenção do timbre médio, ao passo que tônicas não contíguas se mostraram neutras (P. R.: .50).

Tendo em vista a fala de Nova Iguaçu, os índices expressos na Tabela 28 ilustram comportamentos semelhantes aos descritos. Naquela pesquisa, a presença de uma vogal alta tônica contígua à média posterior se destaca como principal contexto de elevação (58%, P. R.: .59).

**Tabela 28** Atuação da variável *Distância entre a pretônica e outra vogal alta na palavra* para o alçamento de /o/, em Nova Iguaçu (cf. ROCHA, 2013: 129)

| VARIÁVEL DISTÂNCIA DE OUTRA VOGAL ALTA |              | [u]     |    |       |
|--|--------------|---------|----|-------|
|  |              | OCO.    | %  | P. R. |
| Tônica                                 | Contígua     | 369/631 | 58 | .59   |
|  | Não contígua | 21/189  | 11 | .37   |
| Átona                                  | Contígua     | 92/674  | 14 | .50   |
|  | Não contígua | 5/89    | 6  | .18   |

c) *Grau de escolaridade dos informantes*

O nível de escolaridade dos informantes foi a terceira variável destacada como relevante para a manifestação da pronúncia [u] no PST. Grosso modo, estudos sobre a elevação das vogais médias em diferentes variedades do Português comprovam que a retenção do timbre médio é mais frequente na maior parte dos dialetos do PB e que pronúncias altas se destacam no PE, independentemente da formação intelectual dos entrevistados. A variedade urbana do Português de São Tomé parece refletir a norma padrão europeia, sendo a frequência da variante [u] diretamente proporcional aos anos de formação dos informantes.

**Tabela 29** Atuação da variável *Escolaridade dos informantes* para o alçamento de /o/, no PST

| VARIÁVEL ESCOLARIDADE DOS INFORMANTES             | [u]       |    |            |
|---|-----------|----|------------|
|   | OCO.      | %  | P. R.      |
| Fundamental – 5 a 8 anos de escolaridade          | 277/754   | 36 | .42        |
| Médio – 9 a 11 anos de escolaridade               | 330/887   | 37 | .44        |
| <b>Superior – mais de 11 anos de escolaridade</b> | 537/1.052 | 51 | <b>.60</b> |

Dito de outra forma, indivíduos com formação básica (P. R.: .42) e média (P. R.: .44) são mais resistentes à aplicabilidade da regra e, nos mais escolarizados, a tendência à pronúncia [u] é maior (P. R.: .60). É notório salientar que as ocorrências foram extraídas de diálogos entre um documentador e um informante. Neles, os entrevistados sabem que estão sendo gravados, apesar de ignorarem os propósitos reais das entrevistas. Trata-se de um contexto razoavelmente monitorado em que a tendência ao alçamento é menor em informantes menos cultos do que naqueles mais escolarizados.

A comparação entre os pesos relativos constatados nesta investigação e aqueles descritos em Rocha (2013) (cf. Tabela 30) sugere a existência de normas padrão distintas nas variedades brasileira e santomense da Língua Portuguesa, ao menos no tocante ao alçamento da média recuada. A variedade brasileira costuma atribuir à pronúncia [o] certo prestígio. Por conseguinte, os percentuais de elevação seriam inversamente proporcionais à escolaridade dos informantes, isto é, maior naqueles menos escolarizados (25%) e menor naqueles com maior formação (17%). Em contrapartida, na variedade urbana do Português de São Tomé, cuja norma de referência é o Português Europeu, o oposto parece se apresentar.

**Tabela 30** Atuação da variável *Escolaridade dos informantes* para o alçamento de /o/, em Nova Iguaçu (cf. ROCHA, 2013: 131) e no PST

| VARIÁVEL <i>ESCOLARIDADE DOS INFORMANTES</i> | [u]            |           |            |           |           |            |
|--|----------------|-----------|------------|-----------|-----------|------------|
|  | NOVA IGUAÇU/RJ |           |            | SÃO TOMÉ  |           |            |
|  | OCO.           | %         | P. R.      | OCO.      | %         | P. R.      |
| Nível Fundamental                            | 356/1.430      | <b>25</b> | <b>.58</b> | 277/754   | 36        | .42        |
| Nível Médio                                  | 225/1.148      | 20        | .49        | 330/887   | 37        | .44        |
| Nível Superior                               | 293/1.670      | 17        | .44        | 537/1.052 | <b>51</b> | <b>.60</b> |

d) *Qualidade da vogal da sílaba seguinte*

A qualidade da vogal da sílaba subsequente foi a quarta variável mais determinante para o alçamento da média posterior. Nessa variável, as ocorrências se distribuem conforme a altura da vogal contígua e vogais altas, sobretudo a não homorgânica (P. R.: .68), favorecem a elevação. Ao mesmo tempo, segmentos baixos (P. R.: .36) e médio-altos (P. R.: .44) inibem a regra, sendo a manutenção mais atuante, precedendo-os.

**Tabela 31** Atuação da variável Tipo de vogal contígua para o alçamento de /o/, no PST

| VARIÁVEL <i>TIPO DE VOGAL CONTÍGUA</i>            |                                | [u]     |      |            |
|---|--------------------------------|---------|------|------------|
|   |                                | OCO.    | %    | P. R.      |
| <b>Vogal tônica/átona (oral/nasal/nasalizada)</b> | <b>Alta homorgânica</b>        | 180/423 | 42.5 | <b>.54</b> |
|   | <b>Alta não homorgânica</b>    | 381/766 | 50   | <b>.68</b> |
|   | Médias-altas                   | 320/725 | 44   | .44        |
|   | Baixas (médias-baixas e baixa) | 263/779 | 34   | .36        |

Além de distribuir os dados de Nova Iguaçu pelos diferentes traços de altura das vogais, Rocha (2013) avalia a influência de segmentos tônicos e átonos e, com base nos índices reproduzidos na Tabela 32, afirma:

embora a vogal [u] acentuada também favoreça a aplicação da regra (.57), os pesos relativos e índices percentuais corroboram a defesa de que [i] seja mais influente no alçamento da pretônica em ambas as séries, já que o alçamento de /o/ também é mais provável em contextos em que, sucedendo a vogal alvo, se encontra a tônica alta não homorgânica (.91) ou um segmento

átono equivalente (.62). Em contrapartida, a manutenção de timbre é mais provável se, após a vogal átona, existe uma tônica não alta, já que o menor peso lhe foi atribuído, quando o fator de referência é o alçamento (ROCHA, 2013: 124).

**Tabela 32** Atuação da variável Tipo de vogal contígua para o alçamento de /o/, em Nova Iguaçu (cf. ROCHA, 2013: 124)

| VARIÁVEL TIPO DE VOGAL CONTÍGUA         |  | [u]       |           |            |
|---|--|-----------|-----------|------------|
|   |  | OCO.      | %         | P. R.      |
| Vogal tônica<br>(oral/nasal/nasalizada) | Alta homorgânica                           | 27/133    | 20        | .57        |
|   | <b>Alta não homorgânica</b>                | 348/521   | <b>67</b> | <b>.91</b> |
|   | Não alta (média baixa, média alta e baixa) | 240/1.992 | 12        | .33        |
| Vogal átona<br>(oral/nasal/nasalizada)  | Alta homorgânica                           | 27/309    | 9         | .43        |
|   | Alta não homorgânica                       | 72/390    | 18        | <b>.62</b> |
|   | Não alta (média baixa, média alta e baixa) | 160/903   | 18        | .51        |

Constata-se, desse modo, a atuação do processo de harmonização vocálica, nas duas amostras e em ambas as séries, sendo os traços de altura assimilados na elevação e na retenção das médias.

*e) Modo de articulação da consoante do ataque da sílaba seguinte*

A fim de analisar a atuação da redução vocálica na amostra, a influência do modo de articulação do segmento introdutório da sílaba subsequente à vogal /o/ foi considerada. A variável se associa aos pontos de articulação precedente e do ataque subsequente à vogal alvo e ao modo de articulação da consoante do ataque da sílaba de núcleo pretônico. Se, na mesma sílaba, segmentos líquidos puderam ser agrupados por apresentarem comportamentos similares, nesta variável, oclusivas e nasais passaram a integrar um mesmo grupo de fatores e manteve-se a distinção entre ele e os demais modos previamente estabelecidos.

Oclusivas e nasais predominaram na amostra, mas, juntamente com as fricativas, se aproximam da neutralidade (P. R.: .51 e .52, respectivamente) quando o fator de referência é a elevação de /o/. A pronúncia [u] é tendência precedendo laterais (P. R.: .56) e a manutenção é mais provável antecedendo vibrantes (P. R.: .23).

**Tabela 33** Atuação da variável *Modo de articulação da consoante do ataque subsequente* para o alçamento de /o/, no PST

| VARIÁVEL MODO DE ARTICULAÇÃO<br>(ATAQUE SUBSEQUENTE) | [u]       |    |            |
|--|-----------|----|------------|
|  | OCO.      | %  | P. R.      |
| Oclusivas e nasais                                   | 759/1.760 | 43 | .51        |
| Fricativas   | 220/551   | 40 | .52        |
| <b>Laterais</b>                                      | 130/220   | 59 | <b>.56</b> |
| Vibrantes  | 34/153    | 22 | .23        |

Em Nova Iguaçu, Rocha (2013: 126) descreve motivações distintas no tocante à atuação da mesma variável. Naquele dialeto, “os maiores índices de elevação correspondem, respectivamente, a segmentos nasais (P. R.: .63 e 37%) e fricativos (P. R.: .55 e 20%). Em contrapartida, a manutenção é mais frequente, em presença de segmentos líquidos (P. R.: .41 e 19%) e oclusivos (P. R.: .43 e 11%)”.

f) *Nasalidade da vogal alvo*

A nasalidade foi a sexta variável influente na pronúncia [u]. Descartadas as ocorrências de ditongação, os dados de vogal dita nasal se limitaram a 176 ocorrências, apresentando-se em 135 delas (77%) a variante [o]. O contexto foi descartado e a análise final concentrou-se nas vogais orais e nasalizadas.

**Tabela 34** Atuação da variável *Nasalidade da vogal alvo* para o alçamento de /o/, no PST

| VARIÁVEL NASALIDADE DA VOGAL PRETÔNICA | [u]       |      |            |
|--|-----------|------|------------|
|  | OCO.      | %    | P. R.      |
| Oral                                   | 818/1.983 | 41.5 | .49        |
| <b>Nasalizada</b>                      | 285/529   | 54   | <b>.54</b> |

Como ilustram os valores expressos na Tabela 34, os pesos relativos se aproximam da neutralidade. No entanto, em vogais nasalizadas, há um número menor de ocorrências e uma maior tendência à elevação (P. R.: .54), e nas orais o oposto se verifica (P. R.: .49).

A Tabela 35 compara os índices percentuais e de ocorrências supracitados àqueles apresentados por Rocha (2013) e revela similaridades entre os resulta-

dos das duas análises, ilustrativas das variedades brasileira e santomense do Português.

**Tabela 35** Atuação da variável *Nasalidade da vogal alvo* para o alçamento de /o/, em Nova Iguaçu (cf. ROCHA, 2013: 136) e no PST

| VARIÁVEL NASALIDADE DA VOGAL<br>PRETÔNICA | [u]            |    |            |           |      |            |
|---|----------------|----|------------|-----------|------|------------|
|   | NOVA IGUAÇU/RJ |    |            | SÃO TOMÉ  |      |            |
|   | OCO.           | %  | P. R.      | OCO.      | %    | P. R.      |
| Oral                                      | 601/3.176      | 19 | .51        | 818/1.983 | 41.5 | .49        |
| Nasal                                     | 90/635         | 14 | .35        | -         | -    | -          |
| Nasalizada                                | 183/437        | 42 | <b>.62</b> | 285/529   | 54   | <b>.54</b> |

Considerando não só segmentos orais e nasalizados, mas também os ditos nasais, Rocha (2013: 135) salienta:

Em contraste com o verificado em /e/, a baixíssima produtividade da variante [u] em contextos de travamento nasal reitera o defendido em Câmara Jr, (1970), Callou; Leite (1986), Battisti (1993), dentre outros que constam da Bibliografia deste estudo. No entanto, não anula a relevância da nasalidade sobre o alçamento, já que os pesos relativos sugerem vogais nasalizadas (.62) como mais suscetíveis à elevação, enquanto segmentos orais se aproximam consideravelmente da neutralidade.

g) *Classe gramatical do vocábulo*

A classe gramatical do vocábulo de que a pretônica faz parte foi o sétimo grupo relevante para a implementação de [u] no dialeto investigado. A variável incluía substantivos, adjetivos, formas finitas e não finitas dos verbos, conectores, numerais, advérbios (em geral e em -mente), formas em -(z)inho (a), pronomes oblíquos “comigo” e “consigo” e, de modo mais específico, os itens “porque”, “portanto”, “bocado(s), bocadinho”, “Tomé, tomense(s) e “você”. Como se esclareceu no início desta seção, por apresentarem comportamentos (semi)categoricos em favor de uma das variantes, todos os itens isolados foram descartados. Ao mesmo tempo, os demais conectores, numerais, advérbios, palavras terminadas em -(z)inho e os pronomes “comigo” e “consigo” foram ínfimos na amostra, justificando a concentração dos dados nas quatro primeiras classes supracitadas. Ademais, por apresentarem comportamentos semelhantes, os dados de adjetivos e formas finitas dos verbos foram reunidos em um único fator, que se contrapôs, na análise final, aos substantivos e formas não finitas dos verbos.

**Tabela 36** Atuação da variável *Classe de vocábulos* para o alçamento de /o/, no PST

| VARIÁVEL CLASSES DE VOCÁBULOS         | [u]       |    |       |
|---------------------------------------|-----------|----|-------|
|                                       | OCO.      | %  | P. R. |
| Substantivos                          | 554/1.225 | 45 | .55   |
| Adjetivos e formas finitas dos verbos | 323/851   | 38 | .46   |
| Verbos formas não finitas             | 233/509   | 46 | .43   |

Em todas as classes listadas, os percentuais apontam o predomínio da manutenção do timbre médio e pesos relativos muito próximos da neutralidade. É evidente, entretanto, que grande parte dos itens em que a vogal média posterior atua são substantivos, o que parece favorecer a maior probabilidade de elevação nesse grupo (P. R.: .55), em contraste com as formas não finitas (P. R.: .43) e finitas dos verbos e os adjetivos (P. R.: .46), que se mostraram mais resistentes à aplicabilidade da regra.

#### b) *Frequência de uso do Forro*

A frequência de uso do Forro se mostrou influente na aplicabilidade da regra de elevação em /o/. Como se enfatizou no tratamento da média anterior, a variável considera três possibilidades de frequência, tendo em vista as auto declarações dos entrevistados. Ela foi postulada na intenção de observar se as variantes das médias pretônicas refletem, de algum modo, o contato entre o crioulo mais falado na região em destaque e o Português.

**Tabela 37** Atuação da variável *Frequência de uso do Forro* para o alçamento de /o/, no PST

| VARIÁVEL FREQUÊNCIA DE USO DO FORRO | [u]       |      |       |
|-------------------------------------|-----------|------|-------|
|                                     | OCO.      | %    | P. R. |
| Pouca ou nenhuma                    | 613/1.254 | 49   | .56   |
| Média                               | 472/1.222 | 38.5 | .47   |
| Alta                                | 59/217    | 27   | .35   |

Grande parte das ocorrências distribui-se entre os indivíduos que alegam pouco/nenhum e médio contato com Forro. À medida que somente um dos dezesseite informantes se insere no último grupo, é natural que essa realidade seja menos produtiva, limitando-se a 217 ocorrências, em que a manutenção é mais provável (P. R.: .35). Apesar de os índices relativos às demais possibilidades se

aproximarem da neutralidade, em informantes cujo contato com o Forro é menor ou nulo, a pronúncia [u] é mais provável (P. R.: .56), enquanto a tendência à manutenção se destaca naqueles cuja frequência é intermediária (P. R.: .47).

i) *Sexo dos informantes*

A variável sexo foi apontada pelo *Goldvarb X* como a nona variável relevante para a aplicabilidade da regra de elevação da média recuada. Estudos variacionistas costumam inseri-la entre as características sociais investigadas sob a alegação de que, de acordo com o valor social atribuído a uma variante, a probabilidade de mudança sonora é maior em homens do que em mulheres, pois elas tenderiam a se adequar com mais facilidade à norma de prestígio. Os percentuais inseridos na Tabela 38 demonstram a maior incidência da manutenção de /o/ em ambos os sexos. No entanto, na fala feminina, a tendência à elevação é maior (P. R.: .56).

**Tabela 38** Atuação da variável *Sexo dos informantes* para o alçamento de /o/, no PST

| SEXO DOS INFORMANTES | [u]       |    |            |
|----------------------|-----------|----|------------|
|                      | OCO.      | %  | P. R.      |
| Masculino            | 660/1.620 | 41 | .46        |
| <b>Feminino</b>      | 484/1.073 | 45 | <b>.56</b> |

Segundo Rocha (2013: 134), em Nova Iguaçu, o oposto se verifica: “mesmo sendo evidente a baixíssima produtividade da variante alta em ambos os sexos (...), o alçamento foi mais frequente no discurso de informantes do sexo masculino (22.4%), sendo esse fator mais sujeito à aplicação da regra (.54)”.

**Tabela 39** Atuação da variável *Sexo dos informantes* para o alçamento de /o/, em Nova Iguaçu (cf. ROCHA, 2013: 135) e no PST

| VARIÁVEL SEXO<br>DOS INFORMANTES | [u]            |    |       |           |    |       |
|----------------------------------|----------------|----|-------|-----------|----|-------|
|                                  | NOVA IGUAÇU/RJ |    |       | SÃO TOMÉ  |    |       |
|                                  | OCO.           | %  | P. R. | OCO.      | %  | P. R. |
| Masculino                        | 450/2.011      | 22 | .54   | 660/1.620 | 41 | .46   |
| Feminino                         | 424/2.237      | 19 | .46   | 484/1.073 | 45 | .56   |

O confronto entre as duas investigações sugere que a pronúncia média é mais prestigiada do que a alta na fala de Nova Iguaçu e na variedade urbana do Portu-

guês de São Tomé, a variante [u] se destaca, sendo, talvez, a manutenção mais estigmatizada.

*j) Ponto de articulação da consoante precedente*

Dialogando com outras abordagens sobre a relação entre os processos de redução vocálica e de elevação de /o/, esta investigação considerou também, os pontos de articulação dos segmentos adjacentes ao núcleo pretônico, sendo o ponto de articulação da consoante do ataque precedente à vogal alvo selecionada como variável influente na realização alta. Como aponta a tabela a seguir, o condicionamento se limitou aos pontos labial, alveolar e velar, já que, em virtude da pouca produtividade, descartaram-se os casos de ataque preenchido por alveopalatais e palatais:

**Tabela 40** Atuação da variável *Ponto de articulação da consoante precedente* para o alçamento de /o/, no PST

| VARIÁVEL PONTO DE ARTICULAÇÃO<br>(CONTEXTO PRECEDENTE) | [u]     |      |            |
|--|---------|------|------------|
|  | OCO.    | %    | P. R.      |
| <b>Labiais</b>   | 387/854 | 45.5 | <b>.52</b> |
| Alveolares   | 274/791 | 34.5 | .44        |
| <b>Velares</b>   | 450/822 | 55   | <b>.54</b> |

Mais uma vez, os pesos relativos expressos na Tabela 40 aproximam-se da neutralidade. Velares (P. R.: .54) e labiais (P. R.: .52) estão, entretanto, mais propensos à elevação do que alveolares (P. R.: .44).

Na tentativa de observar a influência do ponto de articulação da consoante que precede a vogal /o/ em Nova Iguazu, Rocha (2013) distribui as ocorrências em [+ coronais] (alveolares, alveopalatais e palatais) e [-coronais] (labiais e velares). Em meio às peculiaridades de cada investigação, as duas abordagens revelam comportamentos semelhantes no que diz respeito à atuação dessa variável, pois em Rocha (2013: 132), “vogais posteriores precedidas por consoantes [-coronais] se destacaram (...) como o ambiente mais favorável à aplicação da regra (.55), (...) ao passo que segmentos [+coronais] precedendo a vogal alvo parecem motivar sua preservação como média [o]”.

*l) Faixa etária dos informantes*

A faixa etária dos informantes foi a última variável importante para a pronúncia [u] na amostra. Em conformidade com outros estudos variacionistas, ela

foi cogitada na tentativa de atestar indícios de uma mudança em andamento ou de estabilidade entre os processos de alçamento e de manutenção na realidade linguística em pauta.

A Tabela 41 descreve os resultados encontrados na variedade urbana do Português de São Tomé. À semelhança do verificado em /e/, todos os percentuais e ocorrências revelam a maior representatividade da manutenção, com pesos relativos muito próximos da neutralidade. Em indivíduos de meia idade (P. R.: .46), todavia, a manutenção é mais provável, enquanto naqueles com idade superior a 56 anos, a pronúncia [u] se destaca (P. R.: .55) e os mais jovens se mostraram neutros (P. R.: .50).

**Tabela 41** Atuação da variável *Faixa etária dos informantes* para o alçamento de /o/, no PST

| VARIÁVEL FAIXA ETÁRIA    | [u]       |      |            |
|--------------------------|-----------|------|------------|
|                          | OCO.      | %    | P. R.      |
| 25 – 35 anos             | 422/1.001 | 42   | .50        |
| 36 – 55 anos             | 364/884   | 41   | .46        |
| <b>56 anos em diante</b> | 358/804   | 44.5 | <b>.55</b> |

#### 4. CONSIDERAÇÕES FINAIS

Em linhas gerais, a investigação da aplicabilidade da regra de elevação de /e/ e de /o/ aponta baixos índices (input .42 para a média anterior e .33 para a média recuada), demonstrando uma maior tendência à manutenção do timbre médio na fala de São Tomé. Grande parte dos condicionamentos selecionados para a implementação da regra se aproxima da neutralidade e, à medida que a recorrência de um mesmo item se mostrou influente em determinados resultados, acredita-se na interferência do léxico e da coatuação de motivações nos *corpora* pesquisados.

Cumpramos ressaltar o fato de as amostras investigadas unirem ocorrências submetidas ao alçamento por harmonização e por redução vocálicas, o que pode justificar a seleção de uma quantidade superior de variáveis e a sobreposição de gatilhos. No que remete ao primeiro fenômeno, as pronúncias [i] e [u] sobressaem, quando i) na sílaba subsequente à pretônica, há uma vogal alta acentuada [i] (P. R.: .64) e [u] (P. R.: .72); vogais altas (acentuadas ou não) sucedem a pretônica anterior (P. R.: .55) e a vogal [i] é contígua à posterior (P. R.: .68). Constituem indícios da atuação do fenômeno de redução a maior incidência da pronúncia [i], i) antes de consoantes fricativas (P. R.: .58), nasais (P. R.: .56), laterais (P. R.: .54) e [+coronais] (P. R.: .52); ou depois de fricativas (P. R.: .66). Na série recuada, a

elevação é mais provável, se a vogal é *i*) precedida por oclusivas (P. R.: .61), velares (P. R.: .54) e labiais (P. R.: .52); ou sucedida por um ataque preenchido por laterais (P. R.: .56), fricativas (P. R.: .52) e oclusivas e nasais (P. R.: .51). Talvez, a observação isolada de cada processo auxilie o esclarecimento das verdadeiras motivações e dos *inputs* de cada um deles, mas o tempo previsto para o desenvolvimento desta pesquisa não permitiu apurar a questão.

Mesmo constatando a maior influência de motivações estruturais, não se desprezaram os perfis sociais dos entrevistados e o seu contato com o crioulo majoritário na região. As quatro variáveis extralinguísticas foram selecionadas como influentes na elevação de /e/ e de /o/. A fim de comparar sua atuação nas duas amostras, reiteram-se, nas Tabelas 42, 43, 44 e 45, os percentuais de alçamento por série, tendo em vista os fatores que as compõem.

Relativos à faixa etária e ao sexo dos entrevistados, os resultados apresentados nas Tabelas 42 e 43 revelam comportamentos singulares. Todos os percentuais indicam a maior incidência da manutenção nas amostras, mas, se, em /e/, informantes do sexo masculino (37.5%) e da faixa intermediária alçam mais (40.5%); em /o/, a elevação é recorrente no sexo feminino (45%) e em pessoas mais velhas (44.5%).

**Tabela 42** Atuação da variável *Faixa etária dos informantes* para o alçamento de /e/ e de /o/, no PST

| VARIÁVEL FAIXA ETÁRIA | [i]       |             | [u]       |             |
|-----------------------|-----------|-------------|-----------|-------------|
|                       | OCO.      | %           | OCO.      | %           |
| 25 – 35 anos          | 502/1.359 | 37          | 422/1.001 | 42          |
| 36 – 55 anos          | 446/1.107 | <b>40.5</b> | 364/884   | 41          |
| 56 em diante          | 268/868   | 31          | 358/804   | <b>44.5</b> |

**Tabela 43** Atuação da variável *Sexo dos informantes* para o alçamento de /e/ e de /o/, no PST

| VARIÁVEL SEXO DOS INFORMANTES | [i]       |             | [u]       |           |
|-------------------------------|-----------|-------------|-----------|-----------|
|                               | OCO.      | %           | OCO.      | %         |
| Masculino                     | 805/2.144 | <b>37.5</b> | 660/1.620 | 41        |
| Feminino                      | 411/1.190 | 34.5        | 484/1.073 | <b>45</b> |

A Tabela 44 contrasta os índices percentuais e de ocorrências encontrados para a variável Frequência de uso do Forro nas duas séries:

**Tabela 44** Atuação da variável *Frequência de uso do Forro* para o alçamento de /e/ e de /o/, no PST

| VARIÁVEL FREQUÊNCIA DE USO DO FORRO | [i]       |             | [u]       |           |
|-------------------------------------|-----------|-------------|-----------|-----------|
|                                     | OCO.      | %           | OCO.      | %         |
| Pouca ou nenhuma                    | 590/1.536 | <b>38.5</b> | 613/1.254 | <b>49</b> |
| Média                               | 572/1.559 | 36.5        | 472/1.222 | 38.5      |
| Alta                                | 54/239    | 22.5        | 59/217    | 27        |

Nos dois casos, quanto menor o contato com o Forro, maiores os percentuais de elevação, sugerindo alguma relação entre a língua oficial (o Português) e a manifestação do processo.

Por fim, de acordo com a Tabela 45, o número total de ocorrências e os percentuais das pronúncias [i] e [u] são diretamente proporcionais à formação dos entrevistados, sendo mais frequentes quanto mais elevado é o grau de escolaridade.

**Tabela 45** Atuação da variável *Escolaridade dos informantes* para o alçamento de /e/ e de /o/, no PST

| VARIÁVEL ESCOLARIDADE DOS INFORMANTES             | [i]              |           | [u]       |           |
|---|------------------|-----------|-----------|-----------|
|   | OCO.             | %         | OCO.      | %         |
| Fundamental – 5 a 8 anos de escolaridade          | 252/818          | 31        | 277/754   | 36.5      |
| Médio – 9 a 11 anos de escolaridade               | 405/1.212        | 33.5      | 330/887   | 37        |
| <b>Superior – mais de 11 anos de escolaridade</b> | <b>559/1.304</b> | <b>43</b> | 537/1.052 | <b>51</b> |

Ao longo deste capítulo, sempre que possível, compararam-se os comportamentos das vogais médias pretônicas em diferentes variedades do Português, com base em dados contemporâneos de Nova Iguaçu (ROCHA, 2013) e do PST. Naquela investigação, a variável escolaridade foi selecionada apenas na série posterior, mas divulgam-se os índices percentuais e de ocorrências em uma das etapas do estudo. A fim de relacionar, mais uma vez, tais realidades linguísticas, a Tabela 46 retoma os resultados expostos em Rocha (2013: 137) e indica que a elevação

na fala de Nova Iguaçu varia em função da série analisada. Com índices muito baixos e próximos entre si, em /e/, o maior percentual de alçamento coincide com o nível superior de escolaridade. Já, em /o/, o oposto se verifica, uma vez que os percentuais decrescem à medida que a escolaridade aumenta.

**Tabela 46** Atuação da variável *Escolaridade dos informantes* para o alçamento de /e/ e de /o/, na fala de Nova Iguaçu (cf. ROCHA, 2013: 137)

| VARIÁVEL <i>ESCOLARIDADE</i><br><i>DOS INFORMANTES</i> | [i]       |           | [u]       |           |
|--|-----------|-----------|-----------|-----------|
|  | OCO.      | %         | OCO.      | %         |
| Ensino Fundamental                                     | 307/2.003 | 15        | 356/1.430 | <b>25</b> |
| Ensino Médio   | 282/1.941 | 14.5      | 225/1.148 | 20        |
| Ensino Superior  | 357/2.269 | <b>16</b> | 293/1.670 | 17        |

Associadas às evidências do PE descritas na seção 2.1, as semelhanças constatadas permitem que se insira o PST em um *continuum* afrobrasileiro. Em um extremo cujas normas são a elevação e a redução vocálicas se inseriria o Português Europeu, noutro em que a manutenção é a preferência, em meio a casos de alteamento, se situaria o Português do Brasil e, entre ambos, mas mais próximo da realidade brasileira, parece estar localizada a variedade urbana do Português de São Tomé.

**Quadro 6** Esboço de um *continuum* do alteamento de /e/ e de /o/ nas variedades brasileira, europeia e santomense da língua portuguesa

| [+ alteamento] |               | [-alteamento]                                    |  |
|----------------|---------------|--|--|
|                |               |  |  |
| [i] PE         |               | PST (F) = 31%<br>PST (M) = 33,5<br>PST (S) = 43% | PB (F) = 15%<br>PB (M) = 14,5%<br>PB (S) = 16% |
| [u] PE         | PST (S) = 51% | PST (F) = 36,5%<br>PST (M) = 37%                 | PB (F) = 25%<br>PB (M) = 20%<br>PB (S) = 17%   |

**Legenda** – Níveis de escolaridade: (F) = Fundamental / (M) = Médio / (S) = Superior

Variedades do Português: PE: Português Europeu/ PB: Português do Brasil/ PST: Português de São Tomé

## REFERÊNCIAS

- ARAUJO, G. A. Empréstimos recentes do Português, variação fonética e a sílaba na língua são-tomense da ilha de São Tomé. *Papia*, São Paulo, v. 17, p. 55-66, 2007.
- BISOL, L. A. *Harmonia vocálica*: uma regra variável. 1981. Tese (Doutorado em Linguística) – Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 1981.
- BISOL, L. A neutralização das átonas. *Revista DELTA*, São Paulo, v. 19, n. 2, p. 267-276, 2003.
- BORTONI, S. M.; GOMES, C. A.; MALVAR, E. da S. 1992. A variação das vogais médias pretônicas no Português de Brasília: um fenômeno neogramático ou de difusão lexical? *Revista de Estudos da Linguagem*, Belo Horizonte, UFMG, ano 1, v. 1, p. 9-29, 1992.
- BRANDÃO, S. F.; ROCHA, F. M. V.; SANTOS, E. R. Vogais médias pretônicas em início de vocábulo na fala do Rio de Janeiro. *Letras & Letras: estudos em fonologia*, Uberlândia: EDUFU, v. 28, p. 244-257, 2012. 182.
- CALLOU, D.; LEITE, Y. *Iniciação à fonética e à fonologia*. Rio de Janeiro: Jorge Zahar, 2005.
- CÂMARA JR., J. M. *Estrutura da língua portuguesa*. Rio de Janeiro: Vozes, 1970.
- CÂMARA JR., J. M. *Para o estudo da fonêmica portuguesa*. Rio de Janeiro: Padrão, 1977.
- CARMO, M. N. *As vogais médias pretônicas dos verbos na fala culta do interior paulista*. 2009. Dissertação (Mestrado em Linguística) – Faculdade de Letras, Universidade Estadual Paulista, São Paulo, 2009.
- FERRAZ, L. I. *The creole of São Tomé*. Johannesburgo: Witwatersrand University Press, 1979.
- FERRAZ, L. I. Portuguese creoles of West Africa and Asia. In: GILBERT, G. G. (Ed.). *Pidgin and creole languages: essays in memory of John E. Reinecke*. Honolulu: University of Hawai Press, 1987. p. 337-360.
- GONÇALVES, R. M. G. *Propriedades de subcategorização verbal no Português de São Tomé*. 2010. Dissertação (Mestrado em Linguística) – Faculdade de Letras, Universidade de Lisboa, Lisboa, 2010.
- GONÇALVES, R. M. G.; HAGEMEIJER, T. O Português num contexto multilingue: O caso de São Tomé e Príncipe. *Revista Científica da Universidade Eduardo Mondlane*, Moçambique, v. 1, n. 1, p. 87-107, 2015.
- HAGEMEIJER, T. *Clause structure in Santomé*. 2007. Tese (Doutorado em Linguística) – Faculdade de Letras, Universidade de Lisboa, Lisboa, 2007.
- HAGEMEIJER, T. Initial vowel agglutination in the Gulf of Guinea creoles. In: ABOL, E.; SMITH, N. (Eds.). *Complex processes in new languages*. Amsterdam/Filadélfia: John Benjamins Publishing Company, 2009 c. p. 29-50.

- HAGEMEIJER, T. comparação sintáctica. *Papia*, São Paulo, v. 22, n. 2, p. 233-251, 2012.
- MATEUS, M. H. M.; D'ANDRADE, E. *The phonology of portuguese*. Oxford: Oxford University Press, 2000.
- MOTA, J. A.; CARDOSO, S. A. M. Para uma nova divisão dos estudos dialetais brasileiros. In: MOTA, J. A.; CARDOSO, S. A. M. (Eds.). *Documentos 2: Projeto Atlas Lingüístico do Brasil*. Salvador: Quarteto, 2006. p. 15-34.
- NASCENTES, A. *O linguajar carioca*. 2. ed. Rio de Janeiro: Organização Simões, 1953 [1922].
- NASCIMENTO, Fabiane de Mello Vianna da Rocha Teixeira Rodrigues do. O sistema vocálico do Português de São Tomé e o comportamento das vogais médias em contexto pretônico. Tese (Doutorado em Letras Vernáculas/Língua Portuguesa) – Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2018.
- ROCHA, Fabiane de Mello Vianna. da. *O comportamento das vogais médias pretônicas na fala de Nova Iguaçu*. Dissertação (Mestrado em Letras Vernáculas/Língua Portuguesa) – Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2013.
- SILVEIRA, A. C. *Ditongos no Português de São Tomé e Príncipe*. Dissertação (Mestrado em Filologia e Língua Portuguesa) – Faculdade de Letras, Universidade de São Paulo, São Paulo, 2013.



# VOGAIS EM CONTEXTO POSTÔNICO MEDIAL NO PORTUGUÊS DE SÃO TOMÉ

**DANIELLE KELY GOMES**

## **1. INTRODUÇÃO**

Reflexões sobre o vocalismo átono sempre estiveram em pauta nos estudos sobre a Fonologia do Português, dada a variedade de processos que ocorrem nesses contextos. É sabido que o sistema vocálico de sete elementos, pleno na pauta tônica, sofre reduções, em função da atuação de processos fonológicos.

No âmbito do Português do Brasil (daqui por diante PB), uma ampla – porém longe de estar esgotada – descrição do comportamento das vogais nessa variedade pode ser observada em Brandão (2015). Em geral, o sistema vocálico átono se reduz em função do grau de atonicidade da sílaba, coexistindo três subsistemas: o pretônico, com cinco elementos, dada a perda de contraste entre as vogais médias; o postônico final, com três segmentos, em função da neutralização entre as vogais médias e altas; e o postônico não final, a primeira sílaba postônica de um vocábulo proparoxítono, que constitui a grande controvérsia dos estudos sobre o vocalismo no PB, pois – para além de debates sobre o grau de reduções que ocorrem nesse contexto – processos de apagamento são produtivos em diversas normas brasileiras. Uma reflexão mais detalhada sobre esse contexto, no que tange à sua configuração e aos processos de apagamento, será apresentada posteriormente.

No Português Europeu (doravante PE), as descrições revelam que o sistema tônico de sete segmentos<sup>1</sup> apresenta uma reconfiguração para a pauta átona como um todo, sem a distinção entre graus de atonicidade das sílabas. Mateus e D'Andrade (2001, p. 17-18) indicam que o vocalismo átono da variedade se reestrutura em um sistema de quatro elementos, distribuído em dois graus de abertura : três vogais altas ([i], [i̥] e [u]) e uma a que os autores denominam como média [ɐ]. Outro aspecto que marca a diferença entre os contextos átonos das variedades brasileira e europeia é a incidência do processo de apagamento de vogais no sistema átono em geral do PE, nas posições não acentuadas como um todo.

Se há descrições disponíveis e muito consistentes sobre processos fonológicos que atuam no vocalismo átono do PB e do PE, o quadro para as variedades africanas do Português é diametralmente inverso. De uma forma geral, as descrições sobre as variedades africanas do Português e sobre os crioulos de base lexical portuguesa são escassas, e se concentram em aspectos morfossintáticos, mais salientes em situações de contato entre línguas. Aspectos fonológicos dessas variedades há pouco tempo vêm ocupando a agenda de reflexões sobre o “Português em África”.

Este capítulo se insere em uma proposta mais ampla de reflexão sobre aspectos fonético-fonológicos de uma variedade africana do Português, a falada em São Tomé e Príncipe (PST). Mais especificamente, as discussões aqui se concentram na observação do comportamento das vogais postônicas não finais, a primeira vogal postônica de um vocábulo proparoxítono. Conforme dito anteriormente, a pauta acentual desse contexto é controversa nos estudos sobre a Fonologia do Português Brasileiro, e até certo ponto esclarecida no âmbito do PE, dado o fato de que o apagamento de vogais átonas, independentemente da posição que ocupam, é bastante produtivo nessa variedade

Assim, tenta-se observar se é possível propor um continuum afro-brasileiro no comportamento das vogais postônicas não finais. Petter (2007: 9) defende que

as diferentes situações de contato, em épocas diversas, mas envolvendo o português e um conjunto de línguas muito próximas, as do grupo banto, produziram alguns resultados semelhantes nos níveis fonológico, lexical e morfossintático, que nos permitem defender a existência de um *continuum* afro-brasileiro de português. Mesmo considerando que não existam entidades homogêneas identificáveis como “português africano”, “português mo-

<sup>1</sup> Mateus e D'Andrade (2001) defendem que o PE e PB possuem, no subsistema tônico, o mesmo número de segmentos no nível fonético, já que [ə] – categorizado por alguns estudos como um elemento fonológico que contrasta com [a] – na verdade causa um “contraste aparente”, pois “a vogal acentuada [ə] é uma realização alternativa de outras vogais acentuadas em certos contextos” (p. 19).

çambicano”, “português angolano” ou “português brasileiro”, a história do contato e os aspectos linguísticos comuns a essas variedades autorizam-nos a levantar a hipótese desse *continuum*.

Para desenvolver o tema, este capítulo se desdobra por mais seis seções: em (2), apresentam-se as particularidades do contexto postônico não final; em (3), discutem-se aspectos relacionados à caracterização do sistema vocálico do Português de São Tomé e do Forro; nas seções (4) e (5), indicam-se, respectivamente, as hipóteses de trabalho e os fundamentos teórico-metodológicos que norteiam a investigação; em (6) expõem-se os resultados da análise, e em (7) tecem-se as considerações finais.

## 2. O CONTEXTO POSTÔNICO NÃO FINAL E SUAS ESPECIFICIDADES

O sistema vocálico do Português é composto por sete segmentos – /a/, /e/, /o/, /ɛ/, /o/, /i/, /u/ – plenamente especificados na posição tônica. Nas posições não acentuadas, processos de redução são verificados, com o funcionamento de subsistemas que se definem em função da posição da vogal em relação ao acento. As reconfigurações dos subsistemas átonos estão correlacionadas à perda de força expiratória na emissão de segmentos sonoros não acentuados. A perda de força expiratória leva a processos de redução, que afetam as vogais médias.

Câmara Jr. (1970) defende que há uma assimetria na atuação dos processos de redução no contexto postônico não final, com a neutralização entre vogais médias e altas no âmbito da série posterior, mas a manutenção da oposição entre médias e alta na série anterior, redundando em um subsistema de quatro elementos (/a/, /e/, /i/, /u/). Bisol (2003) considera tal assimetria um processo não natural. A autora, com o suporte da Teoria Autossegmental e com base na análise de dados recolhidos no âmbito do projeto VARSUL, defende a tese de que no contexto postônico não final há a flutuação entre os sistemas de três e cinco segmentos, prototípicos das posições átona final e átona pretônica, respectivamente.

Santos (2015), a partir da análise de dados espontâneos e monitorados de *corpora* representativos das normas fluminenses, constata que as vogais médias [e, o] e altas [i, u] estão em variação no Rio de Janeiro, mas que a tendência é que se neutralizem em favor das vogais altas, levando à implementação de um subsistema de três elementos – /a/, /i/, /u/ – semelhante ao verificado na posição átona final. Em seus termos (op. cit, p. 189):

as variantes médias [e o] e altas [i u] estão em variação na fala fluminense, mas fonologicamente verifica-se a redução do quadro para /i a u/. Pode-se concluir que a assimetria entre as vogais /e/ e /o/ postônicas mediais, percebida por Câmara Jr. na fala carioca da década de 1940, embora seja aparen-

temente uma distinção de nível fonológico, na verdade é uma questão de competição entre variantes fonéticas que são mais ou menos produtivas a depender da influência da escolaridade, do perfil idiossincrático do falante e do contexto discursivo.

Já o PE apresenta uma configuração para o sistema átono em geral muito diferente da observada no PB. Mateus e D’Andrade (2000: 134) afirmam que “o português europeu é uma língua em que as vogais não acentuadas são reduzidas e frequentemente apagadas. Isso é mais notável nas posições pós-acental e final”.

Entretanto, além das diferenças em termos de configuração dos quadros vocálicos átonos, PB e PE divergem quanto à frequência da regra de apagamento de vogais. No PE, a não realização de vogais átonas é um processo produtivo, que afeta o vocalismo átono em geral. O PB ainda está “um passo atrás” na implementação de regras de apagamento, já que são mais produtivas as regras de alteamento vocálico. Gomes (2012, 2015, 2018) investiga, sob a perspectiva variacionista, o processo de apagamento das vogais postônicas não finais em dados espontâneos e semidirigidos representativos das normas fluminense e lisboeta. Seus resultados corroboram as constatações de Mateus; D’Andrade, sobre as diferenças quantitativas e qualitativas entre as variedades.

A autora observa que o processo de apagamento das postônicas não finais tem input baixo de aplicação nos dados espontâneos do PB por ela analisados (.07 para as normas culta e popular rural; .04 nos dados da norma popular urbana), e é um fenômeno condicionado por restrições linguísticas e sociais. Os dados espontâneos do PE vão em direção a um input alto de aplicação da regra (.56), condicionada exclusivamente por restrições de ordem linguística.

Uma possível explicação para as diferenças entre as variedades reside justamente no estágio de implementação das regras de alteamento em cada sistema vocálico átono: no PE, o processo de alteamento em contexto pretônico se consolida a partir do século XVIII, com posterior generalização para os demais contextos átonos. Pode-se postular que o apagamento de vogais seja uma etapa posterior à consolidação das regras de alteamento vocálico. O PB, por seu turno, ainda apresenta variação entre as vogais médias, sobretudo no contexto pretônico, um processo que – nesse contexto – é estável. Logo, se entendemos que o apagamento de vogais é uma etapa seguinte à de alteamento, justifica-se a baixa produtividade de apagamento de vogais no PB e a valorização social negativa que o processo recebe em determinados contextos linguísticos e sociais.

O problema que se coloca é onde o PST se situaria tendo em vista as características que distinguem a variedade brasileira da europeia no que concerne ao vocalismo átono? Seria possível, no que concerne ao comportamento das vogais postônicas não finais, estabelecer um *continuum* entre as variedades? Se partirmos

do pressuposto de que o Português de São Tomé toma como norma de referência o Português Europeu, possivelmente se esperaria um alinhamento entre PST e PE na aplicação da regra de apagamento de segmentos vocálicos em geral, com foco específico na sílaba átona não final.

Todavia, é imprescindível observar de que maneira o contato entre línguas em São Tomé molda a norma do Português lá consolidada. Se, por um lado, é possível observar que as consequências da situação de contato entre línguas afetam de tal forma o PST que chega a aproximá-lo do PB no comportamento de algumas variáveis linguísticas, de outro, também é possível notar que o multilinguismo pode direcioná-lo para um caminho que o distancia tanto de sua “norma de referência” quanto da variedade que, por conta de aspectos históricos relacionados à situação de contato entre línguas, com ele mais se pareceria. É o que veremos mais adiante.

### 3. A CARACTERIZAÇÃO DO SISTEMA VOCÁLICO EM SÃO TOMÉ: UMA COMPARAÇÃO ENTRE O FORRO E O PORTUGUÊS

Descrições sobre o sistema vocálico tanto do PST quanto das línguas que com ele coexistem no contexto insular apenas mais recentemente estão em evidência. Para a compreensão dos aspectos que caracterizam a variedade e que possam oferecer subsídios para a percepção da dimensão do contato entre línguas para a formatação do sistema vocálico do PST, é importante retomar as considerações apresentadas em Ferraz (1979), uma descrição sobre o Forro, crioulo de base lexical portuguesa falado por 36,3% dos que habitam a ilha de São Tomé<sup>2</sup>, local onde foram recolhidos os dados desta investigação.

Ferraz (1979, p. 19-20) apresenta, para o Forro, um quadro vocálico em posição tônica idêntico ao verificado no PB, com sete vogais orais e cinco vogais nasais, conforme ilustra o Quadro 1:

**Quadro 1** Vogais do Forro (posição tônica)

|            | ANTERIORES |       |       | CENTRAL | POSTERIORES |       |       |
|------------|------------|-------|-------|---------|-------------|-------|-------|
|            | ALTA       | MÉDIA | BAIXA | BAIXA   | ALTA        | MÉDIA | BAIXA |
| Não nasais | i          | e     | ɛ     | a       | u           | o     | ɔ     |
| Nasais     | ĩ          | ẽ     | –     | ã       | ũ           | õ     | –     |

Fonte: Ferraz (1979, p. 20)

<sup>2</sup> Dados do Censo 2012. Disponíveis em: <<http://www.ine.st/>>. Acesso em: 30 de março de 2018.

O autor não traz apreciações específicas sobre o vocalismo átono do Forro, mas, ao descrever processos fonéticos que afetam os itens lexicais do Português incorporados ao crioulo e a estrutura silábica, tece considerações importantes sobre o comportamento dos segmentos átonos, como se observa nos fragmentos a seguir:

- (i) Palavras portuguesas têm tipicamente um número de sílabas maior do que o das palavras do Forro. As palavras do Forro são, geralmente, dissilábicas. Essa diferença na estrutura silábica resulta em um frequente apagamento de segmentos quando palavras portuguesas são incorporadas ao Forro. Esse apagamento – na forma de aférese, síncope ou apócope – pode envolver vogais átonas, sílabas ou sequências de segmentos. Segmentos acentuados em português não são apagados. (FERRAZ, 1979, p. 38)
- (ii) Por conta de uma tendência derivada do Banto para a estrutura CVCV, palavras portuguesas com mais de duas sílabas são frequentemente reduzidas à forma dissilábica. Em palavras portuguesas com mais de duas sílabas, uma delas com a estrutura CV, em que C é uma líquida, V é geralmente apagada, e a consoante líquida forma um grupo com uma consoante de outra sílaba, usualmente a primeira, como em: Ptg xícara /'ʃikɐrɐ/ -> ST 'şikla; Ptg máscara /'maʃkɐrɐ/ -> ST 'mlaška. (idem, p. 47)

A partir das observações, percebe-se que o Forro apresenta tendência ao apagamento de segmentos átonos em diversos contextos, e à simplificação da estrutura silábica, processos que se verificam em crioulos de uma maneira geral.

Com relação a variedades do Português falado em São Tomé, Silveira (2010, 2013) observa, para o subsistema tônico oral, o mesmo quadro de sete vogais que caracteriza o PB. A diferença entre o que ele chama de Português Vernacular de São Tomé e Príncipe (PVS) e o PE reside no fato de que no PVS há “a ausência categórica das vogais médias centrais [ə] e [ɐ] (entre as vogais tônicas)” (p. 36).

No âmbito da pauta pretônica, Silveira constata o subsistema com cinco segmentos, semelhante ao do PB, com a neutralização de /e, ɛ/ e /o, ɔ/ em favor das médias altas, /e/ e /o/. Nascimento (2018)<sup>3</sup> também investigou o contexto pretônico, com base em dados extraídos do *corpus* Variedades do Português (VA-POR), o mesmo utilizado nesta pesquisa. Seus resultados, contudo, divergem, em alguns aspectos, das tendências apresentadas por Silveira (op. cit).

<sup>3</sup> Para maiores detalhes, sugere-se a leitura do capítulo de Nascimento, nesta obra, sobre o comportamento do contexto pretônico no *corpus* Variedades do Português.

Em ambas as investigações, constata-se que a regra variável de alteamento das vogais médias é um processo presente no PST. Contudo, há diferenças quantitativas entre as duas análises. Enquanto Silveira (2013, p. 39) observa uma mudança em curso, com a realização de /e, o/ como [i, u], com forte tendência para a reconfiguração do sistema de cinco para três segmentos /a, i, u/, Nascimento (2018) nota em seus dados uma situação distinta, com a manutenção de ambas as médias. Para a autora (p. 176),

as vogais médias pretônicas tendem à manutenção do timbre médio, pois todas as rodadas sugeriram baixos índices de aplicação das variantes [i] e [u]. Nesse sentido, pode-se dizer que, embora no padrão de referência europeu o alçamento se destaque, a manutenção parece predominar no dialeto investigado

No que concerne ao contexto átono final, Silveira (op. cit) observa que o sistema do PVS é idêntico ao verificado no PB, com três segmentos /a, i, u/. O mesmo quadro, em suas análises, figura no contexto postônico medial. Os resultados desta investigação, como serão discutidos adiante, divergem da caracterização apresentada pelo autor.

#### 4. HIPÓTESES DE INVESTIGAÇÃO

A partir das considerações tecidas anteriormente, sobre a caracterização do vocalismo do Forro e descrições relativas ao sistema vocálico na variedade vernacular falada em São Tomé, o propósito deste capítulo é investigar como se comportam os dados com relação aos processos de (1) apagamento de vogais no contexto postônico medial e (2) elevação das vogais médias no mesmo contexto. Os dois processos, no âmbito do PB, são variáveis; no que respeita ao PE, o apagamento é variável, enquanto a elevação já se encontra implementada no sistema.

A pergunta que se coloca é: qual seria o alinhamento do PST frente às variedades brasileira e europeia?

Partindo-se da hipótese de que o contato com outras línguas exerceria um papel fundamental na realização de um padrão acental que é não natural até mesmo para indivíduos que possuem português como L1<sup>4</sup> e que não convivem em situação de multilinguismo generalizado, espera-se que índices elevados de apaga-

---

<sup>4</sup> Araújo *et al.* (2007: 37-38) realizaram um levantamento da produtividade dos padrões acen-tuais, com base no levantamento de todos os verbetes do *Dicionário Houaiss da Língua Portuguesa*. Os autores chegaram a um total de 150.875 palavras, das quais 18.413 são proparoxítonas, o que equivale a somente 12% do léxico.

mento da vogal postônica medial sejam encontrados no PST. Com relação ao estatuto das vogais médias no contexto átono medial da variedade, espera-se encontrar variação entre as vogais médias e altas, que se aplicaria tanto na série anterior quanto na série posterior. Parte-se de tal hipótese por conta da variação entre vogais médias e altas no contexto pretônico, e assume-se que o contexto postônico não final também se caracterize por tal competição, mas com dimensões distintas.

## 5. FUNDAMENTAÇÃO METODOLÓGICA

A investigação empreendida neste capítulo tem por suporte teórico-metodológico a Teoria da Variação e Mudança (WEINREICH; LABOV; HERZOG, 1968). Foram analisados 12 inquéritos relativos à localidade de São Tomé, recolhidos no âmbito do projeto *Variedades do Português* (VAPOR), sob o abrigo do Centro de Linguística da Universidade de Lisboa.

No total, levantaram-se 467 dados, analisados com o auxílio do *software* Goldvarb-X. Para a investigação do processo de apagamento da átona não final, postularam-se dez variáveis linguísticas – a natureza dos contextos antecedente e subsequente, os traços de articulação das vogais tônica, postônica medial e postônica final, a estrutura da sílaba postônica não final, a classe morfológica do vocábulo, o número de sílabas da palavra e a produtividade do item no léxico –, e quatro variáveis sociais – sexo, idade, escolaridade e frequência de uso de um crioulo (Brandão, 2011). Na observação do comportamento das vogais médias no mesmo contexto, a análise – por conta do número de dados – é qualitativa, com observação dos índices percentuais e distribuição das ocorrências pelo léxico.

## 6. RESULTADOS

No que se refere ao processo de regularização dos proparoxítonos a paroxítonos a partir da aplicação da regra de apagamento da vogal postônica medial, a distribuição dos dados de apagamento/manutenção da postônica medial no *corpus* analisado está expressa na Tabela 1, a seguir.

**Tabela 1** Distribuição dos dados

|                     | APL | %     | EXEMPLO                   |
|---------------------|-----|-------|---------------------------|
| apagamento da vogal | 162 | 34,7% | <i>século</i> – [ˈsɛklɨ]  |
| manutenção da vogal | 305 | 65,4% | <i>lógica</i> – [ˈlɔʒɨkɐ] |
| total               | 467 | 100%  |                           |

Os resultados revelam que o processo de apagamento da postônica medial é um fenômeno produtivo no PST (34,7%). Por se revelar como uma regra variável, é possível um tratamento estatístico dos dados, em busca dos condicionamentos linguísticos e sociais que se apresentem como relevantes para a ocorrência do fenômeno. No Quadro 2, a seguir, indicam-se as restrições que condicionam a ocorrência do fenômeno em foco.

**Quadro 2** Condicionamentos estatisticamente relevantes para o apagamento da postônica medial no Português de São Tomé (PST)

| Condicionamentos                |              |
|---------------------------------|--------------|
| natureza do contexto precedente |              |
| escolaridade                    |              |
| frequência de uso de um crioulo |              |
| classe morfológica              |              |
| Input                           | inicial: .35 |
|                                 | seleção: .30 |
| Significância                   | .001         |

No Quadro 2, acima, evidencia-se que, das 14 variáveis independentes postuladas, 04 – duas linguísticas e duas sociais – foram consideradas relevantes estatisticamente. A seguir, tecem-se considerações sobre esses condicionamentos. A análise considera, em um primeiro momento, as variáveis linguísticas, e, posteriormente, as restrições sociais.

### 6.1 Variáveis linguísticas: a consoante precedente e a classe morfológica

As análises de cunho variacionista sobre o processo de apagamento da postônica medial em variedades do PB (CAIXETA, 1989; AMARAL, 2000; SILVA, 2006, 2010; FONSECA, 2007; LIMA, 2008, 2017; RAMOS, 2009; CHAVES, 2011; GOMES, 2012) são unânimes em destacar o papel dos segmentos adjacentes à vogal para a aplicação da regra: se há a possibilidade de a consoante que acompanha a vogal ser rressilabificada, a queda do segmento vocálico é favorecida. No âmbito do PE (FERNANDES, 2007; GOMES, 2012), a queda da vogal postônica medial promove a formação tanto de encontros consonânticos que obedecem às condições de boa-formação de sílaba em português quanto de sequências não admitidas fonologicamente, mas possíveis do ponto de vista fonético. Na Tabela 2, a seguir, apresentam-se os índices percentuais e pesos relativos para o contexto precedente à postônica medial.

**Tabela 2** Atuação da variável *Natureza do contexto precedente* para o apagamento da vogal postônica não final

| NATUREZA DO CONTEXTO PRECEDENTE   | EXEMPLO                      | APL/OCO         | P. R |
|-----------------------------------|------------------------------|-----------------|------|
| Consoantes obstruintes            | <i>típico</i> – [ˈtipku]     | 140/274 = 51.1% | .68  |
| Consoantes nasais                 | <i>nômada</i> – [ˈnõmde]     | 11/68 = 14.1%   | .26  |
| Consoantes líquidas               | <i>católica</i> – [kaˈtoʎke] | 8/57 = 13.6%    | .21  |
| Onsets vazios e ataques complexos | <i>fábrica</i> – [ˈfabrkɐ]   | 4/36 = 10%      | .23  |

Nos dados, conforme evidencia a Tabela 2, a presença de uma consoante obstruinte no ataque da sílaba postônica medial é, de acordo com os índices estatísticos, o fator preponderante para a implementação do apagamento da vogal átona medial (.68). Todos os demais contextos fônicos considerados, por estarem muito abaixo do ponto neutro, não foram considerados relevantes do ponto de vista estatístico.

Todavia, uma compreensão mais abrangente do papel do contexto só é possível quando se considera também a *Natureza do segmento subsequente* à postônica medial, dada a pressão exercida pela possibilidade de ressilabificação das consoantes adjacentes quando ocorre a queda da vogal. Ferraz (1979, p. 47), destaca que no Forro a síncope de vogais é favorecida, nos vocábulos com mais de duas sílabas, quando há a presença de consoante líquida em uma estrutura silábica CV, o que forma um ataque complexo e, por consequência, diminui a palavra para que o item de origem portuguesa se adapte ao padrão dissilábico, prototípico do crioulo. Na Tabela 3, a seguir, apresentam-se índices numéricos para os segmentos no contexto subsequente à vogal postônica medial. Apesar de não ter sido indicado como relevante<sup>5</sup>, vale aqui uma apreciação dos resultados para esse condicionamento.

**Tabela 3** Natureza do contexto subsequente

| NATUREZA DO CONTEXTO       | EXEMPLO                     | APL/OCO         | P. R  |
|----------------------------|-----------------------------|-----------------|-------|
| Consoantes obstruintes     | <i>hipótese</i> – [iˈpõtsɪ] | 106/296 = 35,8% | (.44) |
| Consoantes nasais          | <i>mínimo</i> – [ˈmĩmɯ]     | 46/121 = 38%    | (.60) |
| Consoante líquida lateral  | <i>oráculo</i> – [oˈraklɯ]  | 4/8 = 50%       | (.75) |
| Consoante líquida vibrante | <i>chácara</i> – [ʃakrɐ]    | 7/20 = 25,9%    | (.58) |

<sup>5</sup> Neste caso, os índices estatísticos foram retirados da primeira rodada do step-down.

Na tabela acima, os índices estatísticos relativos à atuação das variantes consoante líquida lateral, nasal e líquida vibrante indicam serem esses contextos os que mais favorecem a aplicação da regra de apagamento (.75, .60 e .58, respectivamente), sendo as obstruintes as consoantes que bloqueiam a aplicação da regra. De certa forma, observa-se que a mesma tendência que se verifica, conforme Ferraz (op. cit.), no Forro se mantém no PST, também se observa nas variedades do PB: quando a queda da postônica medial promove a formação de *cluster* complexo formado por uma consoante obstruinte e uma líquida, o fenômeno é amplamente favorecido. No PE, conforme citado anteriormente, a condição de boa-formação pode ser violada, promovendo, no nível fonético, o encontro de consoantes potencialmente não ressilabificáveis.

Com relação à atuação da outra variável linguística selecionada, *classe morfológica do vocábulo*, os resultados expressos na Tabela 4 revelam que os nomes são os itens atingidos pela regra de apagamento, enquanto os verbos se mostram como contextos morfológicos de resistência.

**Tabela 4** Atuação da variável *Classe morfológica do vocábulo para o apagamento da vogal postônica não final*

| CLASSE | EXEMPLO                          | APL/OCO         | P. R |
|--------|----------------------------------|-----------------|------|
| Nomes  | <i>doméstica</i> – [du'mɛʃtkɐ]   | 159/412 = 38,6% | .53  |
| Verbos | <i>ajudávamos</i> – [aʒu'davmuʃ] | 6/55 = 10,9%    | .27  |

Uma possível justificativa para esses índices pode residir no fato de que, nos verbos, a postônica medial se localiza nas desinências de tempo e modo (dos pretéritos imperfeito e mais-que-perfeito do indicativo e no pretérito imperfeito do subjuntivo). Por estar no contexto de informação morfológica, talvez a retenção da vogal átona medial seja o mais prototípico nessa classe morfológica. Entretanto, o número pequeno de dados inviabiliza a investigação mais profunda dessa hipótese.

## 6.2 Variáveis sociais: escolaridade e frequência de uso de um crioulo

Do ponto de vista dos condicionamentos extralinguísticos, o programa de análises estatísticas apontou que as variáveis *Escolaridade* e *Frequência de uso de um crioulo* são restrições que condicionam o apagamento da postônica medial nos dados do PST.

Os resultados para as duas variáveis têm de ser observados em função da relação que se estabelece entre os condicionamentos. Na Tabela 5, a seguir, estão expressos os índices para a variável *escolaridade*.

**Tabela 5** Atuação da variável *Escolaridade* para o apagamento da vogal postônica não final

| ESCOLARIDADE      | APL/OCO        | P.R |
|-------------------|----------------|-----|
| Nível Fundamental | 8/75 = 10.7%   | .14 |
| Nível Médio       | 65/154 = 42.2% | .64 |
| Nível Superior    | 89/238 = 37.4% | .60 |

Os resultados da tabela acima revelam serem os indivíduos de nível médio de instrução aqueles que mais favorecem o apagamento da postônica (.64), seguidos dos de nível superior (.60). Os indivíduos de nível fundamental de instrução atuam como bloqueadores do processo (.14). Tal resultado soa contraditório, já que as análises sobre o fenômeno em outras variedades do Português indicam serem os indivíduos de escolaridade baixa os que mais favorecem a aplicação da regra. Contudo, um olhar atento sobre a variável justifica os índices encontrados.

Para a apreciação dos índices relativos à escolaridade, é necessário que se reflita sobre o papel da *Frequência de uso de um crioulo*, indicado na Tabela 6, a seguir<sup>6</sup>. Vale destacar que esta variável originalmente é controlada em três níveis distintos de uso de crioulos (frequência baixa, frequência média e frequência alta). Contudo, nesta investigação, não foram levantados os dados nos inquiridos relativos aos informantes que se expressam regularmente em crioulo.

**Tabela 6** Atuação da variável *Frequência de uso de um crioulo* para o apagamento da vogal postônica não final

|                  | APL/OCO         | P. R |
|------------------|-----------------|------|
| Frequência baixa | 108/323 = 33,4% | .43  |
| Frequência média | 54/144 = 37.4%  | .63  |

Não é possível tratar a escolaridade sem considerar o grau de contato que esses informantes mantêm com os crioulos que coexistem com o PST. Na Tabela 6, acima, observa-se que os indivíduos que se reconhecem como falantes eventuais de

<sup>6</sup> A variável Frequência de uso de um crioulo foi formulada por Brandão (2016, p. 91) nos seguintes termos: “frequência (a) zero/baixa, referente aos indivíduos que se expressam fundamentalmente em português; (b) média, relativa aos indivíduos que se expressam em português, mas dominam um crioulo e dele fazem uso eventualmente; (c) alta, abarcando os indivíduos que, embora falem o português e o tenham como L1, se expressam, regularmente, num crioulo.”

crioulo tendem a apagar a postônica medial (.63), enquanto os que se reconhecem como falantes apenas de português são menos suscetíveis ao processo (.43).

Quando se cruzam as variáveis sociais selecionadas, é possível justificar o motivo pelo qual os falantes do nível intermediário são aqueles que mais implementam o apagamento da postônica medial. Os resultados estão indicados na Tabela 7, a seguir.

**Tabela 7** Cruzamento entre as variáveis *Escolaridade* e *Frequência de Uso de um crioulo*

| NÍVEL       | FREQUÊNCIA BAIXA | FREQUÊNCIA MÉDIA |
|-------------|------------------|------------------|
| Fundamental | 5/8 = 62.5%      | 3/8 = 37.5%      |
| Médio       | 14/65 = 21.5%    | 51/65 = 78.4%    |
| Superior    | 89/89 = 100%     | –                |

Os indivíduos que se reconhecem como falantes eventuais de crioulo se concentram, nesse conjunto de dados, entre os de nível médio de escolarização. Os índices expressos na Tabela 7, acima, reforçam a tendência de que os falantes de Forro, por conta de esse crioulo, em geral, regularizar palavras polissilábicas portuguesas em dissílabos, apliquem com mais frequência a regra de apagamento da vogal átona medial (78,4%) do que os informantes que se reconhecem como falantes apenas de Português, no nível médio de escolarização (21,5%).

Sobre o processo de elevação das vogais médias na pauta postônica medial, antes da apresentação dos resultados verificados na amostra, é importante retomar algumas considerações sobre o contexto.

Santos (2015), em uma análise que busca discutir o estatuto fonológico das vogais átonas não finais no PB e no PE, destaca que, na variedade europeia, estão consolidadas:

- (i) as regras de alteamento, tanto na série anterior como na série posterior;
- (ii) a simultânea posteriorização de /e/;
- (iii) a elisão de vogais.

Já na variedade brasileira, seus resultados apontam para:

- (i) os momentos finais de implementação do alteamento de /e/ e /o/;
- (ii) o apagamento de vogais, mas em proporções menores do que as verificadas no PE;
- (iii) variação entre as articulações médias e altas das vogais, processo condicionado por aspectos sociais e estilísticos.

De Paula; Brandão (2012), com base na análise de dados de *corpora* socio-linguísticos representativos da fala culta (NURC) e da fala popular (PEUL e APERJ) do Rio de Janeiro, constatam que a vogal média posterior vem se mostrando mais propensa ao alteamento, enquanto a anterior tende a manter a realização como média na maior parte dos dialetos, mesmo não estando imune ao alçamento. As autoras observam que (p. 147)

a depender da variedade, pode-se ter, em contexto postônico não final, um quadro fonológico simétrico – de cinco (/i e a o u/) ou três (/i a u/) vogais, ou um quadro assimétrico de quatro vogais (/i e a u/), como o que se verifica na variedade culta carioca (...).

No Português de São Tomé, Silveira (2013) atesta que o quadro postônico não final é configurado em uma série de três elementos, [i, a, u], e que o Português Vernacular de São Tomé difere do PE por não apresentar o schwa [ə], sendo o /a/ realizado como [a]. O PVS também se diferenciaria do PB pois, para Silveira, na variedade brasileira, o quadro postônico medial teria a mesma configuração do contexto pretônico<sup>7</sup>, enquanto que, no PVS, os quadros postônicos não final e final seriam idênticos, sem a possibilidade de manutenção da realização média para /e/ e /o/.

Os dados recolhidos no *corpus* Variedades do Português diferem dos resultados verificados por Silveira em sua amostra. Mesmo com o pouco número de ocorrências de palavras que apresentam uma vogal média na sílaba postônica medial, o que impede uma análise quantitativa refinada, os índices expressos na Tabela 8, a seguir, vão em direção a um quadro muito semelhante ao verificado por Câmara Jr. e por De Paula; Brandão na norma culta carioca.

**Tabela 8** Realização das médias postônicas não finais

|     | REALIZAÇÃO COMO MÉDIA | ELEVAÇÃO  | TOTAL DE DADOS |
|-----|-----------------------|-----------|----------------|
| /e/ | 17 = 80%              | 4 = 20%   | 21             |
| /o/ | 0                     | 16 = 100% | 16             |

Os resultados expressos na Tabela 8 sugerem, para o PST, a mesma assimetria proposta por Câmara Jr.: variação no âmbito da série anterior, com a compe-

<sup>7</sup> Nas palavras do autor (op. cit., p. 40):

“No sistema do PE descrito por Mateus e D’Andrade (2000), é apresentado o conjunto dos seguintes fones na posição postônica não final: [i, a, u], enquanto que no PB (variedade utilizada por Câmara Jr., 1995 [1970]), é apresentado o seguinte conjunto [i, u, e, o, a]”.

tição entre as realizações [e] e [i] para a /e/; e alteamento na série posterior, com a neutralização entre média e alta no contexto, com a realização de /o/ como [u]. Um olhar detalhado para os itens lexicais encontrados no *corpus*, expressos na Tabela 9, a seguir, permite uma análise mais abrangente dos processos.

**Tabela 9** Itens lexicais proparoxítonos com vogal média na sílaba postônica medial e suas realizações fonéticas

| /e/  |   | /o/  |
|--|---|--|
| [e]  | [i]   | [u]  |
| gên[e]ro (8 ocorrências)<br>fenôm[e]no (3 ocorrências)<br>núm[e]ro (2 ocorrências)<br>út[e]ro<br>cér[e]bro<br>cél[e]bre<br>alfând[e]ga | indíg[i]nas<br>quilôm[i]tro<br>voltáss[i]mos<br>líd[i]res | ép[o]ca (11 ocorrências)<br>incôm[o]do (2 ocorrências)<br>monót[o]na<br>mêt[o]do<br>Espín[o]la |
| 17/21 = 80%  | 4/21 = 20%  | 16/16 = 100%   |

Na Tabela 9, acima, fica patente o número reduzido de itens com vogais médias postônicas na amostra (11 itens distintos com vogal /e/, 5 itens diferentes com vogal /o/), o que, evidentemente, nos permite apenas sugerir que, na amostra, o panorama que se verifica é o mesmo que vigorava na norma carioca culta descrita em Câmara Jr. (1970)

Evidentemente, a não naturalidade do quadro assimétrico tende a ser resolvida. E um possível caminho para essa resolução talvez possa redundar na estabilização do contexto postônico medial em um quadro com três elementos. Contudo, somente uma investigação mais consistente, com um maior número de dados e com o controle de outros contextos além dos inquéritos do tipo DID, pode trazer uma luz sobre a configuração do sistema vocálico postônico medial da variedade.

## 7. CONSIDERAÇÕES FINAIS

As análises realizadas neste trabalho permitem as seguintes observações quanto aos dois processos investigados.

No que se refere ao apagamento da vogal postônica não final, os resultados sugerem que a redução dos vocábulos proparoxítonos ao padrão acentual paroxítono é um processo bastante frequente (34.7%), condicionado por fatores lingüís-

ticos (*natureza do segmento precedente e classe morfológica do vocábulo*) e sociais (*escolaridade e frequência de uso de um crioulo*).

Já em relação ao comportamento das vogais médias postônicas não finais, os indícios revelam que o PST – no conjunto de dados analisados – apresenta uma configuração muito semelhante à observada por Câmara Jr. para a norma culta carioca, com variação entre a vogal média e a alta no âmbito da série anterior (/e/ – [e] ~[i]), em favor da manutenção da vogal média e implementação da regra de alteamento na série posterior (/o/ – [u]).

Com relação à proposição de um *continuum* afro-brasileiro para os processos, talvez não seja possível – nesse conjunto de dados – afirmar que, em ambos os fenômenos investigados, o PST estaria se aproximando do PB ou do PE, este último, em teoria sua “norma de referência”.

Conforme demonstrado ao longo do capítulo, as regras de apagamento de vogais no PB ainda estão sujeitas a valoração negativa, e os inputs baixos de ocorrência do processo podem corroborar a estigmatização do apagamento. No PE, as alterações profundas no vocalismo átono levam, como etapa posterior, à implementação de regras de apagamento de segmentos vocálicos. Já o PST, em função da influência do crioulo Forro, tende muito mais a aplicar processos de elisão de segmentos, em uma tentativa de regularizar os itens com mais de duas sílabas ao padrão dissilábico, não estando as regras de apagamento vinculadas a um espelhamento no PE.

No que concerne às regras de alteamento das vogais médias postônicas mediais, os dados aqui apresentados revelam que o PST estaria em um estágio anterior ao que o PB (a norma fluminense) se encontra hoje. Todavia, os poucos dados e a falta de um controle em outros contextos impedem afirmações mais consistentes sobre o comportamento de /e/ e /o/ postônicos não finais na variedade. Está aí uma agenda que urge ser cumprida para uma melhor caracterização do vocalismo átono não final do Português de São Tomé.

## REFERÊNCIAS

AMARAL, Marisa Porto do. *As proparoxítonas: teoria e variação*. Porto Alegre, PUC-RS, 2000. 235.fls. Tese (Doutorado em Letras), Pontifícia Universidade Católica do Rio Grande do Sul.

ARAÚJO, Gabriel Antunes; GUIMARÃES-FILHO, Zwinglio O.; OLIVEIRA, Leonardo; VIÁRIO, Mario. As proparoxítonas e o sistema acentual do português. In: ARAÚJO, Gabriel Antunes de *et al* (Org.). *O acento em português: abordagens fonológicas*. São Paulo: Parábola, 2007. p. 37-60.

BRANDÃO, Sílvia Figueiredo. O cancelamento da marca de número nominal na variedade urbana não *standard* do Português de São Tomé. Comunicação apresentada ao XVI

Congreso Internacional de la Asociación de Linguística y Filología de la América Latina (ALFAL). Alcalá de Henares, Madrid, 6-9 de junho, 2011.

BRANDÃO, Silvia Figueiredo. *Variação e o estatuto de variedades do Português. Diadorim*, v. 18, número especial, p. 83-104, 2016.

CAIXETA, Valmir. *Descrição e análise da redução das palavras proparoxítonas*. 1989. Dissertação (Mestrado em Linguística). Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 1989.

CÂMARA JR., Joaquim Mattoso. *Estrutura da Língua Portuguesa*. Petrópolis: Vozes, 1970.

CHAVES. Raquel Gomes. *A redução de proparoxítonos na fala do Sul do Brasil*. 2011. 173f. Dissertação (Mestrado em Linguística). Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, 2011.

DE PAULA, Alessandra; BRANDÃO, Silvia Figueiredo. A assimetria do quadro vocálico em contexto postônico não final: aspectos diacrônicos e realidades sincrônicas. *Signum: estudos da linguagem*. n. 15/1, p. 129-149, jun. 2012.

FERNANDES, Ana Catarina Garcia. *Apagamento de vogais átonas em trissílabos proparoxítonos: um contributo para a compreensão da supressão vocálica em português europeu*. 2007. 155 f. Dissertação (Mestrado em Linguística). Universidade do Porto, Porto, 2007.

FERRAZ, Luiz Ivens. *The creole of São Tomé*. Johannesburg: Witwatersrand University Press, 1979.

FONSECA, Simone Meckler. *O problema das proparoxítonas: a perda da vogal postônica*. 2007. 68 f. Dissertação (Mestrado em Linguística). Universidade Federal de Juiz de Fora, Juiz de Fora, 2007.

GOMES, Danielle Kely. *Síncope em proparoxítonas: um estudo contrastivo entre o português brasileiro e o português europeu*. Rio de Janeiro: UFRJ, 2012. 273 f. Tese (Doutorado em Letras Vernáculas). Faculdade de Letras, Universidade Federal do Rio de Janeiro, 2012.

GOMES, Danielle Kely. O apagamento das vogais postônicas não finais: uma análise contrastiva entre variedades do português. *Revista da ABRALIN*, v. 14: 185-106, 2015.

GOMES, Danielle Kely. *Vocalismo postônico não final no Português de São Tomé*. Trabalho apresentado no Annual Meeting of the Association of Portuguese and Spanish Lexified Lexically-based Creoles. Stockholm: Stockholm University, 13-15 June 2017.

GOMES, Danielle Kely. Síncope das vogais postônicas não finais: uma análise contrastiva entre variedades do português. In: DE PAULA, Alessandra; GOMES, Danielle Kely; SILVEIRA, Eliete; MACHADO VIEIRA, Marcia dos Santos; VIEIRA, Silvia Rodrigues. (Org.). *Uma história de investigação sobre a Língua Portuguesa: homenagem a Silvia Brandão*. São Paulo, Blucher, 2018. p. 213-224.

HAGEMEIJER, Tjerk. As ilhas de Babel: a crioulação no Golfo da Guiné. *Revista Camões*, 6: 74-88, 1999.

HAGEMEIJER, Tjerk. As línguas de São Tomé e Príncipe. *Revista de Crioulos de Base Lexical Portuguesa e Espanhola* (1)1: 1-27, 2009.

LIMA, Giselly de Oliveira. *O efeito da síncope em proparoxítonas: uma análise fonológica e variacionista com dados do sudoeste goiano*. 2008. 195 f. Dissertação (Mestrado em Linguística). Universidade Federal de Uberlândia, Uberlândia, 2008.

LIMA, Giselly de Oliveira. *Percepção da síncope em palavras proparoxítonas*. 2017. 154f. Tese (Doutorado em Linguística). Universidade Federal de Uberlândia, Uberlândia, 2017.

NASCIMENTO, Fabiane de Mello Vianna da Rocha Teixeira Rodrigues do. *O sistema vocálico do português de São Tomé e o comportamento das vogais médias pretônicas*. 2018. 194 f. Tese (Doutorado em Língua Portuguesa). Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2018.

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

PETTER, Margarida Maria Taddoni. Ampliando a investigação do continuum afro-brasileiro do português. *Papia: Revista Brasileira de Estudos do Contato Linguístico*, 25(2): 305-317, 2015.

RAMOS, Adriana Perpétua. *Descrição das vogais postônicas não finais na variedade do Noroeste Paulista*. 2009. 175f. Dissertação (Mestrado em Linguística). Universidade Estadual Paulista, São José do Rio Preto, 2009.

SANTOS, Alessandra de Paula. *Variação e mudança no vocalismo postônico medial em português*. 2015. 200f. Tese (Doutorado em Língua Portuguesa). Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2015.

SILVA, André Pedro. *Supressão da vogal átona postônica não final: uma tendência das proparoxítonas na língua portuguesa com evidências no falar sapeense*. 2006. 139 f. Dissertação (Mestrado em Língua Portuguesa). Universidade Federal da Paraíba, João Pessoa, 2006.

SILVA, André Pedro. *Vogais postônicas não finais: do sistema ao uso*. João Pessoa, 2010, 216f. Tese (Doutorado em Linguística). Universidade Federal da Paraíba, João Pessoa, 2010.

SILVEIRA, Alfredo Christofolletti. *O sistema vocálico do português vernacular de São Tomé e Príncipe. Relatório final (Iniciação Científica)*. Processo FAPESP 2010/07263-9, 2010

SILVEIRA, Alfredo Christofolletti. *Ditongos no Português de São Tomé*. 2013. 109f. Dissertação (Mestrado em Filologia e Língua Portuguesa). Universidade de São Paulo, São Paulo, 2013.

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

# O DITONGO /ei/ NA FALA DE SÃO TOMÉ

RAPHAELA RIBEIRO PASSOS

## 1. INTRODUÇÃO

Pesquisas sobre variedades do Português falado como Língua 1 ou Língua 2 em diferentes áreas da África têm demonstrado haver uma série de convergências entre elas e o Português do Brasil (PB), o que levou alguns linguistas a considerarem a existência de um *continuum afro-brasileiro* (PETTER, 2007; AVELAR; GALVES; 2014), conforme já explicitado na Apresentação deste livro, e que vem sendo confirmado em diversos estudos sobre a concordância nominal e a verbal, entre os quais Brandão; Vieira (2012), Brandão (2013), Vieira; Bazenga (2013), que focalizam o Português de São Tomé (PST), em contraste com o PB.

Neste estudo<sup>1</sup>, realizado com base nos pressupostos teórico metodológicos da Teoria da Variação e Mudança (WEINREICH; LABOV; HERZOG, 1968), pretende-se observar o comportamento do ditongo /ei/ em contexto medial e final de vocábulo na variedade urbana do Português de São Tomé, isto é, se esse diton-

---

<sup>1</sup> Este texto apresenta, na íntegra, alguns trechos da dissertação (Passos, 2018) que deu origem ao presente capítulo.

go se mantém como [ej] – como em af[ej]ção –, se sofre monotongação – [e], como em solt[e]ra e fal[e] – ou, ainda, se concretiza como [ɐj], como em d[ɐj]xa, a exemplo do que ocorre no Português Europeu (PE), que serve de norma de referência ao PST.

Com base nos resultados obtidos, objetiva-se, ainda, verificar não só se, nesse particular, há convergências entre as variedades santomense e brasileira, mas também se determinadas ocorrências se devem a uma possível influência do Forro, o crioulo mais falado em São Tomé. Para tanto, parte-se de amostras selecionadas de entrevistas com falantes da comunidade urbana de São Tomé, extraídas do *Corpus* VAPOR, do Centro de Linguística da Universidade de Lisboa, e realizadas por Tjerk Hagemeyer em 2009.

Para o desenvolvimento deste estudo, fazem-se observações sobre os ditongos em Português (seção 2), seguidas, na seção 3, por comentários sobre alguns estudos que abordam o ditongo /ei/ no PB (3.1) e no PST (3.2). Na quarta seção, apresentam-se os procedimentos metodológicos que nortearam o trabalho e na quinta, expõem-se os resultados da análise da monotongação de /ei/ no PST em contexto medial (5.1) e final de vocábulo (5.2). Na sexta seção, apresentam-se as considerações finais.

## 2. OBSERVAÇÕES SOBRE OS DITONGOS DECRESCENTES EM PORTUGUÊS

Para tratar da questão, selecionaram-se estudos de Câmara Jr. (2011 [1970]), Bisol (1989, 1994, 2012) e Mateus; d'Andrade (2000), embora também tenham focalizado a questão Callou; Leite (2009 [1990]), Gonçalves; Costa (1995), Gonçalves (1997) e Silva (2014).

Segundo a definição de Câmara Jr. (2011[1970], p. 46), o ditongo ocorre “quando a vogal, em vez de ser o centro da sílaba, fica numa de suas duas margens, como as consoantes. O resultado é uma vogal modificada por outra na mesma sílaba e constitui-se o que se chama o ditongo”. Nessa posição, há forte neutralização e o quadro do sistema vocálico se resume apenas à oposição entre a vogal alta posterior e a anterior, como nos exemplos pai/pau e sei/seu.

Câmara Jr. levanta ainda o questionamento sobre o fato de a vogal assilábica ser, na verdade, uma consoante no Português, devido ao funcionamento similar entre esses dois elementos. Ao autor parece melhor considerar tal segmento como vogal, argumentando que, após ditongos, é possível encontrar o /t/ brando, que, em português, só existe depois de vogais (*dinheiro, costureira, Europa* etc.).

Ao tratar a estrutura da sílaba, Câmara Jr. volta a descrever mais detalhadamente os ditongos. Partindo do conceito de sílaba como “*um movimento de as-*

*censão, ou crescente, culminando num ápice (o centro silábico) e seguido de um movimento de decrescente, quer se trate do efeito auditivo, da força expiratória ou da tensão muscular*” (p. 53), o autor conclui que, por suas características, a vogal é normalmente o centro da sílaba em português. Assim, “a estrutura da sílaba depende desse centro, ou ápice, e do possível aparecimento da fase crescente, ou da fase decrescente, ou de uma e outra em volta dele, ou seja, nas suas margens ou encostas” (p. 53).

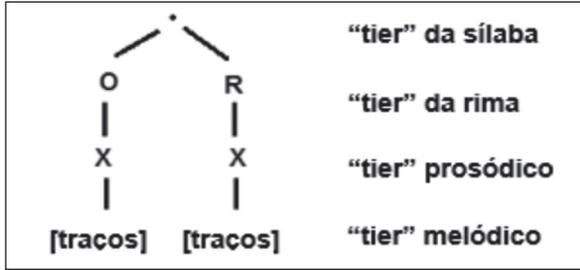
A proposta de representação para o centro da sílaba é V, e para os elementos marginais, C, resultando em alguns possíveis padrões silábicos para o português. De acordo com as possíveis organizações silábicas, surge o questionamento de como classificar a sílaba composta por um ditongo.

Os denominados alofones assilábicos, ou seja, as vogais altas /i/ e /u/ dos ditongos decrescentes, funcionam na sílaba como a parte decrescente, assim como a consoante /l/ e os arquifonemas /R/, /S/ e /N/. Entretanto, embora funcionem como uma consoante, possuem natureza vocálica, dando então origem ao questionamento anteriormente mencionado sobre sua classificação, se a sílaba com ditongo seria do tipo CVC ou do tipo CVV. A principal diferença entre as duas representações é que a primeira se refere a uma sílaba travada, enquanto a segunda, a uma sílaba livre com núcleo ramificado.

Câmara Jr. apresenta três argumentos em defesa da segunda representação: (i) o processo de monotongação, amplamente difundido em português; (ii) a variação livre na divisão silábica no contexto vogal seguida por vogal alta em sílaba átona, como em *vaidade* (*vai-da-de ~ va-i-da-de*); (iii) a possível produção de um /i/ assilábico como /e/ e um /u/ assilábico como /o/ (como em *papaê*). Interpretando, pois, a vogal assilábica como V, surge o problema de distinguir o número de sílabas do vocábulo com ditongo, como, por exemplo, no caso de “peito”. Como determinar que se trata de um dissílabo e não de um trissílabo é uma questão que se apresenta e para a qual Câmara Jr. propõe como solução o caráter de emissão reduzida da vogal assilábica, o que a caracterizaria como uma semivogal.

Além desse problema, há ainda a questão sobre a real existência de ditongos em português, pois há a dúvida se essas sequências não poderiam ser consideradas como hiato, constituído por uma sequência de duas vogais silábicas. A descrição comum gramatical não é satisfatória para Câmara Jr., pois opõe pares como “sai” e “saí”, que não podem ser considerados pares opositivos mínimos já que não possuem a mesma sequência acentual. O autor apresenta, então, alguns exemplos em que a sequência acentual é a mesma, como o par “riu” e “rio” (/riu/ e /ri<sup>u</sup>/) na fala do Rio de Janeiro, o que demonstra, com a oposição, portanto, a existência do ditongo. Com base no exposto, Câmara Jr. enumera 11 ditongos decrescentes, entre os quais /ei/.

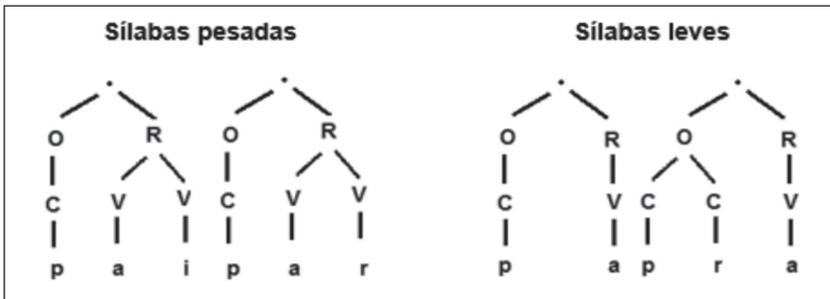
Bisol (1989) apresenta uma análise sobre a estrutura dos ditongos e seus diferentes tipos, que, mais tarde, seria por ela revisitado (1994; 2012). Nesse estudo, em que analisa os ditongos com base na teoria da sílaba, a autora define sílaba “como um objeto multidimensional de sequência de segmentos, cujos constituintes são organizados hierarquicamente” (BISOL, 1989, p. 186). Na Figura 1 é apresentada a mencionada organização hierárquica.



**Figura 1** Organização hierárquica da sílaba (BISOL,1989. p. 186)

O que é definido como *tier* na figura acima são sequências de unidades. Nesse diagrama, é no *tier* prosódico, representado por X, que está o chamado espaço temporal ao qual correspondem os elementos da sílaba, como uma consoante seguida por uma vogal, por exemplo. Acima desse nível, temos o *onset* e a rima da sílaba, que se ligam aos constituintes prosódicos que estão no nível hierárquico abaixo (as consoantes e vogais da sílaba)

Outro ponto importante é a distinção entre sílabas leves e pesadas, sendo as sílabas leves compostas por rima simples, de um só segmento, e as pesadas compostas por rima ramificada, com mais de um segmento. Os respectivos diagramas são apresentados na Figura 2.



**Figura 2** Diferença entre sílaba pesada e sílaba leve (BISOL,1989. p. 187)

A partir de tais apontamentos e de outros princípios norteadores de suas análises, Bisol levanta a discussão sobre dois tipos de ditongos: o **verdadeiro** e o **falso**. Nas palavras da autora, a distinção se dá da seguinte maneira: “o ditongo

pesado, o verdadeiro, [é] associado a duas posições no “tier” da rima, e o ditongo leve, associado a uma só posição. O primeiro constitui uma sílaba complexa e tende a ser preservado; o segundo constitui uma rima simples e tende a ser perdido” (BISOL, 1989, p. 189-190).

Alguns exemplos apresentados para ilustrar o ditongo verdadeiro, ou pesado, são “p[aw]ta” e “r[ej]no”, palavras nas quais o ditongo tende a ser preservado e no âmbito do qual dificilmente ocorreria o fenômeno da monotongação. Já exemplos de palavras com ditongos falsos, ou leves, seriam “p[ej]xe” e “f[ej]ra”, vocábulos cujo ditongo tende a ser perdido, na fala, com alta frequência.

Com base nas observações de Bisol, pode-se considerar o processo de monotongação do ditongo [ej], aqui estudado, como um fenômeno, de fato, passível de ocorrer e presente na língua, devido aos fatores estruturais que o compõem, principalmente no caso dos chamados falsos ditongos. Além disso, Bisol também apresenta observações fundamentais para a definição de variáveis a serem aplicadas para estudos sobre a queda desses ditongos, quando analisa os contextos em que o falso ditongo pode estar inserido. Um desses contextos é aquele em que, no segmento subsequente ao ditongo, se encontra uma consoante palatal. Segundo Bisol, o acréscimo ou apagamento da semivogal diante de uma consoante palatal pode ser feito sem que haja qualquer prejuízo para o significado da palavra. Em suas palavras: “o glide é sempre consequência da palatal. Sua forte relação com a consoante palatal indica que ele se forma no tier melódico, o qual, por sua vez, também é um componente organizado hierarquicamente, com traços ligados a um centro comum, o tier da raiz” (BISOL, 1989, p. 191).

Assim, nesse contexto, o ditongo seria, na verdade, um espriamento do traço [+alto] da palatal subsequente, não presente, portanto, na estrutura subjacente. Outro contexto que criaria um falso ditongo seria quando há a presença de um tepe no segmento subsequente. Nesse caso, como no anterior, a perda da semivogal não leva à criação de pares mínimos e o sentido da palavra é mantido.

Bisol apresenta duas hipóteses para justificar que o ditongo seguido por tepe é, na verdade, um falso ditongo: uma seria por metátese, com espriamento de traços de outros segmentos, gerando a semivogal, e a outra por inserção, considerando a escala de sonoridade do tepe, que, dentre as líquidas, consoantes que possuem certas características vocálicas, seria o segmento mais próximo da escala de sonoridade das vogais.

Entretanto, ambas as hipóteses apresentam problemas e não são capazes de explicar de todo como se daria a inserção da semivogal nesse contexto. Como as hipóteses apresentadas não dão conta de oferecer uma justificativa completa, Bisol apresenta mais um dado para reforçar a defesa de que, diante de tepe, há na verdade um falso ditongo. Esse dado é uma comparação do processo de monoton-

gação dos ditongos decrescentes diante de consoante palatal e diante de tepe. Nos dois contextos, os índices de apagamento da semivogal são altos, atingindo percentuais de 76% e 81%, respectivamente, com pesos relativos de .89 em ambos os casos (BISOL, 1989, p. 197). A alta frequência da monotongação leva a autora a concluir que, assim como quando seguida por palatal, a semivogal do ditongo diante de tepe não está presente na estrutura subjacente na sílaba.

Sobre a oposição entre ditongos verdadeiros e ditongos falsos, a autora conclui com a seguinte sistematização: “ditongos formados no nível mais abstrato, tier da rima, tendem a ser preservados; ditongos mais próximos à superfície, formados no tier melódico, podem ser perdidos” (BISOL, 1989, p. 221).

Bisol revisita o tema posteriormente, em dois momentos: em 1994, em estudo no qual, baseada em análises estatísticas, apresenta mais argumentos para defender a existência do ditongo fonético em oposição ao fonológico, tratando, em especial, do falso ditongo diante de consoante palatal; e, em 2012, focalizando o ditongo diante de tepe.

No primeiro desses trabalhos, são apresentados resultados da análise de uma amostra que contempla sete informantes da cidade de Porto Alegre, mais especificamente os resultados para a variável *contexto subsequente ao ditongo*, variável que interessava ao estudo e que estava constituída pelos fatores consoante palatal e tepe em oposição aos demais fatores (labial, velar, alveolar e vogal). Os resultados apresentados indicam claramente que a queda da semivogal é extremamente favorecida quando há, no contexto seguinte, a presença de uma palatal ou do tepe (BISOL, 1994, p. 124).

Partindo desses resultados, ela volta a defender a ideia de que, nesses contextos específicos, o ditongo não está presente na estrutura subjacente da sílaba. Duas observações importantes são feitas: (i) apenas os ditongos [aj] e [ej] que ocorrem diante de consoante palatal na sílaba seguinte podem sofrer monotongação, independentemente da tonicidade da sílaba; (ii) há palavras que, na fala, apresentam o ditongo, embora este não conste de sua representação gráfica (“fechar”, alternando entre f[e]char e f[ej]char). Segundo Bisol:

Não há como interpretar inserção do glide de um lado e apagamento do outro, em se tratando de contexto similar. Essa é a evidência mais forte que nos leva a admitir que a forma subjacente das palavras [com e sem ditongo diante de consoante palatal na sílaba seguinte] não possuem a vogal alta responsável pelo glide. São palavras de estrutura subjacente de uma vogal só, nessa posição. (1994 p. 126, 127).

Assim, a teoria de que o glide, nesse contexto, surge como um espriamento do traço vocálico da consoante palatal, que é uma consoante complexa, é refor-

çada e mais uma vez defendida por Bisol. A autora conclui que, se o glide que aparece no contexto é um glide derivado, então o ditongo formado também é um ditongo derivado, ou seja, um falso ditongo.

BISOL (2012) é um adendo ao trabalho anterior, e trata do segundo contexto que foi introduzido, mas não analisado, de geração de um falso ditongo com a presença de um tepe como segmento subsequente ao ditongo. Para a análise, Bisol parte das características das consoantes líquidas, ou seja, do fato de estas serem segmentos classificados como aproximantes que possuem o traço vocálico e que constituiriam, juntamente com as vogais, uma única classe, como aproximantes ou como segmentos vocálicos.

Dessa forma, assim como no caso de uma palatal no contexto seguinte, com um tepe nessa posição a formação do ditongo se daria por espraçamento do traço vocálico do tepe. Algumas evidências são apresentadas para sustentar essa hipótese, como o fato de que, no PB, depois de vogal, a líquida lateral (normalmente realizada como velar ou dorsal) geralmente se apresenta como glide posterior [w] (me[w], por *mel*), mas diante do /S/ plural, o glide passa a coronal, em consonância com o traço coronal de /S/, como em m[ej]s. Outra evidência seria encontrada ao se observarem outras línguas, nas quais as líquidas são realizadas como glide, como, por exemplo, no cibaño, um dialeto do espanhol (*carta –caita*). Por fim, há ainda o que se observa no processo de aquisição de linguagem pelas crianças, que costumam realizar as líquidas como uma vogal alta. Todas essas evidências reforçam a ideia central de que o ditongo, em contexto de tepe subsequente, é resultado do espraçamento do traço vocálico do tepe, e, portanto, um falso ditongo.

Cabe, ainda, mencionar, tendo em vista os objetivos deste estudo, como o ditongo se comporta no PE.

Mateus; d’Andrade (2000), na descrição do sistema fonológico do PE, focalizam a estrutura da sílaba, baseados no padrão culto de Lisboa, e observam que o PB e o PE possuem praticamente o mesmo quadro de sete vogais tônicas, como no PB. No PE, além dessas vogais, há ainda o [ɐ], cujo contraste com [a] “*seria apenas aparente, uma vez que [ɐ] tônico é uma realização alternativa de outras vogais tônicas em determinados contextos*” (p. 19): (i) antes de consoante palatal (t[ɐ]lha – “telha”); (ii) antes de glide palatal (l[ɐ]i – “lei”); (iii) antes de consoante nasal (c[ɐ]ma – “cama”).

Devido à possibilidade de ocorrência de [ɐ] antes de glide palatal, há no PE um ditongo que não ocorre no PB, o [ɐj]. Essa é uma das poucas diferenças entre as duas variedades no quadro de ditongos decrescentes tônicos apresentado por Mateus; D’Andrade (2000), que oferecem, como exemplo, as realizações de “queixa” no PE – [kɐj-ʃɐ] – e no PB – [kej-ʃɐ]. A outra diferença é que, no PE, não ocorreria o ditongo decrescente [ɔw], que, no PB, resulta da realização, em final de sílaba, da lateral como um glide (*volta*).

Sobre o fenômeno da monotongação dos ditongos /ei/ e /ai/ no PE, Cunha (1986) afirma ser esse um processo antigo, já registrado pela grafia desde *Os Lusíadas*, com os exemplos “peixe” e “baxa”, sendo temerário, pois, atribuir esse fenômeno a inovações brasileiras. Sobre o ditongo /ou/, o autor afirma ser sua monotongação comum no PE desde o século XVII, não sendo também o fenômeno geral apenas no Brasil.

Cintra (1983) volta-se para a questão dos ditongos /ei/ e /ou/ e sua variação com a forma monotongada e, com base na distribuição dessas variantes, constitui duas isófonas utilizadas por ele para delimitar as áreas linguísticas portuguesas. Segundo sua divisão, as possíveis realizações da forma ditongada /ou/ são características dos dialetos setentrionais, já a variante monotongada tipifica os dialetos centro-meridionais. No caso de /ei/, a variante ditongada se apresenta nos dialetos setentrionais e em uma subárea dos dialetos centro-meridionais, e a variante monotongada nos dialetos centro-interior e sul.

Em suma, no PB e no PE, há a ocorrência de ditongos que podem sofrer monotongação, gerando duas variantes possíveis. Constitui-se como um dos objetivos deste trabalho investigar se no Português de São Tomé tal processo também ocorre e que fatores o condicionam.

### **3. ESTUDOS SOBRE O DITONGO /ei/**

#### **3.1 No Português do Brasil**

No âmbito do PB, diversos foram os estudos que focalizaram a monotongação de /ei/. Dentre eles, citam-se os de Ribeiro (1990) referente ao Sudoeste do Paraná, que observou apenas a atuação de fatores estruturais; o de Paiva (1996), sobre a fala do Rio de Janeiro, retomado em 2011; Lopes (2002), realizado em Altamira, Pará, em que, diferentemente dos demais estudos, uma das variáveis relevantes para a atuação da regra foi o nível de escolaridade do falante; Pereira (2004), que verificou o apagamento de semivogais em ditongos orais na Região Sul do país, mais especificamente, em Tubarão, Santa Catarina; Toledo (2011) que tratou especificamente da monotongação de /ei/ na fala de Porto Alegre, com o objetivo de verificar que contextos linguísticos a condicionam, observando, em especial, a possível atuação de fatores de natureza morfológica.

Tendo em vista a convergência de resultados neles observados, comentam-se, nesta seção, os estudos de Paiva (1996, 2011), apresentando-se, ao final, um quadro comparativo dos resultados obtidos pelos demais autores aqui citados, independentemente das especificidades de cada análise.

Paiva (1996) analisou os ditongos /ou/ e /ei/ com base em entrevistas de fala espontânea realizadas com moradores do Rio de Janeiro, distribuídos por sexo, idade e faixa etária, estabelecendo assim as três variáveis sociais que seriam controladas. Além delas, verificou a atuação de quatro variáveis estruturais: *extensão da palavra*, *contexto subsequente*, *tonicidade da sílaba* e *posição no vocábulo (sufixo ou radical)*. A posição da sílaba no vocábulo, medial ou final, não foi levada em conta, pois, em seu entendimento, em posição final, o ditongo /ei/ sempre seria mantido e /ou/ sempre seria monotongado.

Sua hipótese de partida era que a monotongação de /ei/ é condicionada principalmente por fatores linguísticos. A autora levantou um total de 3133 dados, sendo 2111 do ditongo [ej]. Os resultados encontrados para o ditongo indicaram que, conforme a hipótese inicial, apenas fatores estruturais estariam atuando sobre o fenômeno, já que nenhuma variável social foi selecionada na análise realizada: mostraram-se atuantes o *contexto subsequente*, a *extensão da palavra* e a *posição no vocábulo*.

Para a variável *contexto subsequente*, de acordo com o ponto de articulação, os fatores que se mostraram favorecedores da monotongação foram as velares (P. R.: .89), embora a autora ressalte que a alta frequência do vocábulo “manteiga” deveria ser o gerador desse resultado, já que, em outros vocábulos, o mesmo não ocorreu. Em seguida, vêm e as alveolopalatais (P. R.: .93). Já os demais fatores, dentais, alveolares e vogais, se mostraram inibidores do processo, com pesos relativos de .15, .27 e .10, respectivamente.

Considerando os fatores referentes ao modo de articulação, o tepe se mostrou altamente favorecedor do processo (P. R.: .99), enquanto as fricativas, embora também favorecedoras, apresentaram um peso relativo mais baixo, .56. Os demais fatores, oclusivas, nasais e laterais, se mostraram inibidores da monotongação, com pesos relativos de .13, .13 e .25, respectivamente.

Quanto à *extensão da palavra*, verificou-se que, quanto maior o número de sílabas, maiores as chances de a monotongação ocorrer. Tanto os fatores polissilábicas quanto trissilábicas apresentaram um P. R. alto, de .71, seguido pelas dissilábicas (P. R.: .67). Por fim, as palavras monossilábicas se mostraram altamente inibidoras da monotongação (P. R.: .07).

A última variável selecionada, *posição no vocábulo*, demonstrou que ditongos que ocorrem no sufixo são mais sensíveis ao processo de monotongação (P. R.: .61), ao contrário dos que aparecem no radical, cujo peso relativo (.38) demonstrou que esse contexto é inibidor do processo. Com relação a essa variável, Paiva observa que os resultados podem estar sofrendo interferência de outra variável, o *contexto subsequente*, uma vez que o sufixo derivacional *-eiro* se mostrou bastante presente nos dados, ou seja, a presença de um tepe após o ditongo poderia

estar gerando a queda da semivogal, e não, necessariamente, o fato de a posição ser a de um sufixo.

Em suma, o trabalho de Paiva confirma a hipótese inicial da autora, que constatou que a monotongação do ditongo /ei/ depende basicamente de restrições de ordem estrutural, como o modo e o ponto de articulação do segmento subsequente, sendo favorecido, nesse contexto, pelo tepe e as fricativas, dentre estas as palatoalveolares, e que fatores sociais não estariam interferindo no processo.

Posteriormente, em 2011, Paiva revisita o comportamento do /ei/ no Rio de Janeiro. No estudo, pretendia verificar a atuação da variável *tempo* sobre o processo, associando análises de tempo aparente e de tempo real em um estudo combinado dos tipos painel e tendência. Esse tipo de análise se deve, segundo a autora, ao fato de que “tendências depreendidas em um estudo do tipo painel são reforçadas pelo estudo da comunidade de fala, observando-se notável sistematicidade e regularidade na trajetória do processo” (p. 32). Para tanto, indivíduos entrevistados no final da década de 1990 foram recontactados para novas entrevistas.

Os resultados encontrados com a comparação das duas amostras parecem sugerir um recuo do uso da variante monotongada. Paiva analisa separadamente os 31 indivíduos que apresentaram um incremento nos anos de escolarização no intervalo entre as duas entrevistas e constata que, na maioria desses casos, houve um declínio tanto em frequência quanto em peso relativo da variante monotongada. Entretanto, ao analisar separadamente também o grupo dos indivíduos que não tiveram alteração nos anos de escolaridade, os resultados parecem sugerir que o recuo da variante monotongada pode não estar relacionado diretamente à escolaridade, já que, nesse grupo, também houve um declínio da queda da semivogal: quatro falantes mostraram nítido recuo de uso da variante monotongada; três apresentaram estabilidade no processo e dois deles um aumento de aplicação do processo.

Paiva destaca que uma possível explicação para essas tendências contraditórias pode ser a influência do item lexical mais recorrente em uma entrevista, pois uma análise mais detida da segunda entrevista desses falantes mostra que a variante monotongada está concentrada em itens lexicais em que a semivogal anterior antecede o tepe, como em *dinheiro* e *feira*, contexto em que a realização da variante monotongada se torna praticamente categórica (p. 35).

Considerando tais resultados comparativos entre as duas amostras e os resultados complementares encontrados para indivíduos e para a comunidade, Paiva levanta a hipótese de que haja uma mudança geracional, “com indivíduos mudando ao longo da sua vida e a comunidade se mantendo estável” (p. 37). Assim, os falantes estariam alterando seu comportamento linguístico de maneira a “adequar-se à nova faixa etária” em que se encontram, o que pode ser observado no estudo de tendência, que compara as duas amostras e mostra uma estabilidade do

processo de monotongação nas duas épocas distintas, sendo a faixa 1 a que mais monotonga e a faixa 3 a que menos monotonga.

Com relação aos condicionantes estruturais, Paiva argumenta que sua importância não pode ser ignorada para a determinação da direcionalidade da alternância [ej]/[e]. O ponto e o modo de articulação seguem sendo os principais condicionantes do processo, sendo o tepe e as fricativas palatais os principais favorecedores da queda da semivogal. Entretanto, a autora observa um avanço do processo no ambiente tepe, o que parece constituir uma mudança já consumada, enquanto para o ambiente de fricativa palatal parece haver um recuo da implementação do processo, demonstrando ser esse um contexto de maior variabilidade na realização da semivogal.

Como se observou no início desta seção, apresenta-se, a seguir, um quadro comparativo dos fatores que se mostraram relevantes para a monotongação de /ei/ nos estudos aqui mencionadas, lembrando que a escolha e a composição das variáveis não obedeceram necessariamente aos mesmos parâmetros.

**Quadro 1** Comparação entre os fatores condicionantes da monotongação de /ei/ em seis estudos no âmbito do PB

| Ribeiro (1990)<br>Sudoeste do Paraná                               | Paiva (1996)<br>Rio de Janeiro-RJ  | Lopes (2002)<br>Altamira- PA  | Pereira (2004)<br>Tubarão, SC  | Toledo (2011)<br>Porto Alegre-RS   | Paiva (2011)<br>Rio de Janeiro-RJ  |
|--|--|---|--|--|--|
| <u>Tipo de sílaba:</u> não finais                                  | <u>Contexto subsequente:</u> alvéolopalatais; velares;<br>- tepe;<br>- fricativas      | <u>Localização do ditongo na estrutura morfológica da palavra:</u> base | <u>Monotongação categórica:</u><br>Ressaltando que a grande maioria dos dados era composta por palavras com tepe no contexto subsequente | <u>Contexto seguinte:</u><br>- tepe<br><br><u>Natureza morfológica:</u><br>radical<br><br><u>Classe de palavras:</u><br>não-verbos | Tendência ao recuo da variante monotongada.<br><br>Indícios de mudança geracional, talvez vinculada ao incremento da escolaridade.<br><br>Constante atuação de fatores estruturais: tepe e fricativa palatal no contexto subsequente |
| <u>Número de sílabas:</u><br>três/quatro.                          |  |   |  |  |  |
| <u>Estrutura da sílaba:</u> travadas por /S/                       | <u>Extensão da palavra:</u><br>- polissilábicas;<br>- trissilábicas;<br>- dissilábicas | <u>Contexto fonético seguinte:</u><br>tepe;<br>consoantes palatais      |  |  |  |
| <u>Classe do vocábulo:</u><br>Substantivos/<br>adjetivos           | <u>Posição no vocábulo:</u><br>- sufixo  | <u>Natureza da origem do vocábulo:</u><br>-comum                        |  |  |  |
| <u>Contexto fonológico subsequente:</u><br>tepe;<br>palatoalveolar |  | <u>Escolaridade:</u><br>baixa   |  |  |  |

Fonte: Passos (2018)

### 3.2 No Português de São Tomé

No PB, há diversos estudos sobre o ditongo /ei/, tanto que, na seção anterior, selecionaram-se apenas alguns deles, de modo a retratar a monotongação em diferentes épocas e áreas do país.

A situação no que se refere ao PST é bastante distinta. Embora haja alguns estudos no âmbito morfossintático sobre essa variedade, raros são os que recobrem aspectos fonético-fonológicos, conforme observam Brandão; De Paula no capítulo 4 deste livro. Dentre eles, encontra-se o de Silveira (2013), que procede a uma descrição dos ditongos que ocorrem no PST, entre eles /ei/, também focalizado por Passos; Barcelos (2015), no âmbito da Iniciação Científica da UFRJ.

Silveira (2013) buscou investigar, no português vernacular são-tomense, a realização dos ditongos orais e se eles são passíveis de monotongação, como ocorre, por exemplo, no PB, verificando, também, que fatores a condicionariam. O autor parte da hipótese de que, em São Tomé, as realizações dos ditongos não serão as mesmas que no PE, uma vez que, nessa região, concorrem com o português outras línguas crioulas, que não têm ditongos, ou os têm de forma limitada.

Para o trabalho variacionista, realizado com o auxílio do programa Goldvarb 2001, o autor usou um *corpus* composto por entrevistas de fala espontânea, realizadas entre os anos de 2009 e 2011, totalizando 50 horas de gravações. Na seleção dos informantes, distribuídos por sexo, faixa etária e escolaridade, deu preferência àqueles que fossem naturais de São Tomé e que tivessem pais são-tomenses

Foram controladas, além das variáveis sociais já citadas, as variáveis estruturais *classe morfológica do item lexical, posição do ditongo no vocábulo, tonicidade, localização do ditongo na estrutura morfológica da palavra, contexto fonético subsequente e número de sílabas*.

O total de ocorrências de ditongos orais encontradas foi de 3017, quase todos observados, com exceção de [ɔj], [ɔw], [ɛw] e [ɛj], não produzidos por nenhum dos informantes. Analisando-os em conjunto, Silveira constatou que em 49% dos casos (1482 em 3017) houve monotongação. A maioria deles apresentou algum percentual de monotongação, com exceção de /iw/, /uj/, com apenas três ocorrências no *corpus*: “cuidar”, “viu” e “desistiu”.

Após a análise em conjunto, o autor focalizou cada um dos ditongos. O quadro a seguir mostra os resultados encontrados para os diferentes ditongos estudados (além dos ditongos [iw], [uj], explicados anteriormente).

**Tabela 1** Percentuais de monotongação dos ditongos no PST (SILVEIRA, 2013)

| DITONGO | TOTAL DE OCORRÊNCIAS | PERCENTUAL DE MONOTONGAÇÃO |
|---------|----------------------|----------------------------|
| [ɔj]    | 606/3017             | 35%                        |
| [ow]    | 588/3017             | 99,5%                      |
| [aj]    | 469/3017             | 26%                        |

|      |          |     |
|------|----------|-----|
| [ew] | 387/3017 | 5%  |
| [aw] | 79/3017  | 11% |
| [ej] | 525/815  | 65% |

Fonte: Silveira (2013), com adaptações

Para o ditongo /ei/ foram encontrados 815 dados. Desse total, 65% apresentaram monotongação, o que demonstra ser esse um processo produtivo também nessa variedade. Os resultados indicaram seis variáveis como as mais atuantes para a implementação da monotongação.

Os resultados relativos à variável *contexto seguinte*, a mais saliente delas, indicaram que os fatores que mais favorecem o processo são, em sequência de relevância, os róticos, com P. R.: .96; as fricativas pós-alveolares (P. R.: .93), as fricativas alveolares, (P. R.: .65); e, com um P. R. neutro, .50, as oclusivas alveolares, na realidade, mais especificamente a oclusiva alveolar surda [t]. Esse último fator chama a atenção pelo fato de que, no geral, em outras variedades do Português, ele se mostrar um inibidor do processo, não havendo relato de casos de monotongação do ditongo /ei/ diante de [t].

Quanto à variável *classe de palavra*, há desfavorecimento da monotongação quando a palavra é um verbo, com P. R.: .32. São os não verbos que favoreceriam a regra, com P. R.: .80. Silveira observa que, do total de verbos encontrados (275), apenas em 64 houve a monotongação, e que esses casos não eram com o ditongo em posição final, mas sim medial, como, por exemplo, nos vocábulos “aproveita”, “aleija” e “deixa”. Com o ditongo em posição final, como em “falei”, não foi encontrado nenhum caso de monotongação. Já dentre as palavras classificadas como não-verbos (460), somente em três delas o ditongo estava em posição final.

Tais observações levam à conclusão de que, embora os resultados pareçam indicar que o fator não-verbo favoreça a queda da semivogal e o fator verbo a desfavoreça, parece ser a posição do ditongo no vocábulo o que realmente está condicionando o processo, tanto que a outra variável selecionada é justamente a *posição da sílaba no vocábulo* (inicial, medial ou final). Se o ditongo se encontra em contexto medial, é alto o índice de monotongação de /ei/ (P. R.: .81), diferentemente do que ocorre quando está em posição inicial (P. R.: .51, considerado neutra) ou final (P. R.: .03, desfavorecedor).

A última variável estrutural selecionada como relevante, o *número de sílabas*, demonstrou que o grupo das dissílabas é o maior favorecer da monotongação, com P. R.: .62, seguido pelas polissílabas com P. R. próximo da neutralidade, .53.

As duas últimas variáveis selecionadas foram as de natureza social *escolaridade* e *idade*. Os resultados demonstraram que quanto maior a escolaridade do indivíduo, menor a probabilidade de que ele realize a monotongação. O P. R. do fator *escolaridade alta* foi de .27. O grupo dos falantes de escolaridade média apresentou P. R. próximo da neutralidade, .48. Por fim, os menos escolarizados favorecem a monotongação (P. R.: .64).

Já para a variável *idade*, os resultados indicam que quanto mais jovens forem os falantes, maior a probabilidade de que haja o processo de monotongação. Silveira atribui esses resultados ao fato de os mais jovens serem, geralmente, mais inovadores no uso da língua, enquanto os mais velhos seriam mais conservadores.

Passos; Barcelos (2015) também estudaram esse ditongo, chegando a resultados semelhantes aos encontrados por Silveira. O estudo demonstrou que a variável linguística de maior destaque para a implementação da monotongação é o contexto subsequente, com destaque para as fricativas pós-alveolares e o tepe. Além disso, também foi observada a monotongação diante da dental [t]. Uma divergência encontrada com relação aos resultados de Silveira foi a ocorrência de monotongação de /ei/ em posição final em verbos. Sobre as variáveis sociais, assim como no trabalho exposto anteriormente, escolaridade se mostrou saliente, indicando que, quanto maior a escolaridade do falante, menor a queda da semivogal. Em suma, podem-se observar convergências nos resultados encontrados nos estudos no âmbito do PB e do PST. Em ambas as variedades, os maiores condicionantes do processo de queda do ditongo são os fatores estruturais, principalmente a presença, no contexto seguinte, de um tepe ou de uma fricativa pós-alveolar. A grande diferença entre as duas variedades é que, no PST, a presença de um [t] no contexto seguinte também favorece o processo, enquanto no PB esse fator bloqueia a queda da semivogal. Com relação às variáveis sociais, apenas escolaridade parece ser relevante, com resultados similares nas duas variedades, que apontam que quanto maior a escolaridade do indivíduo, menores as chances de monotongação.

#### 4. ASPECTOS TEÓRICO-METODOLÓGICOS

A realização da pesquisa se deu com base nos pressupostos teórico-metodológicos da Teoria da Variação e Mudança (Weinreich, Labov; Herzog, 2006 [1968]) e da Sociolinguística Variacionista desenvolvida por Labov (2008 [1972]; 1994; 2001), que considera, entre outros princípios, a noção de regra variável, dependente da atuação de restrições sociais e linguísticas.

Serviu de base para o levantamento das ocorrências do ditongo /ei/ o *Corpus VAPOR*, do Centro de Linguística da Universidade de Lisboa, que conta com entrevistas realizadas por Tjerk Hagemeijer em 2009. Todos os informantes, nativos

de São Tomé, declararam ser falantes de Português como L1 e estão distribuídos por sexo, três faixas etárias (18 a 35, 36 a 55 e 56 a 75 anos) e três níveis de escolaridade (fundamental: 5 a 8 anos; médio: 9 a 11 anos e superior: de 12 anos em diante). As gravações, que têm duração média de 25,6 minutos, são do tipo DID (Diálogo entre Informante e Documentador) e versam sobre temas diversos.

Definiram-se, para a análise, realizada com o auxílio do programa estatístico Goldvarb-X, as variáveis clássicas *sexo*, *escolaridade* e *faixa etária*, além da variável comentada anteriormente, *frequência de uso de um crioulo* (BRANDÃO, 2011), aplicada para averiguar se os indivíduos que fazem uso de uma língua crioula apresentam influências dessa língua ao falarem português. Essa variável está dividida em três fatores: (i) frequência *zero ou baixa*, para indivíduos que se expressam fundamentalmente em Português; (ii) frequência *média*, relativa a indivíduos que se expressam em Português, mas têm o domínio de uma língua crioula e a usam eventualmente; e (iii) frequência *alta*, para indivíduos que falam Português, mas se expressam com maior regularidade com um crioulo.

Com a aplicação dessa variável, buscou-se confirmar a hipótese de que o uso regular do Forro gera influências no Português com relação à produção do ditongo /ei/, uma vez que, segundo Ferraz (1979), o padrão silábico principal do Forro é do tipo CV, seguido dos padrões V e CVC, que são menos usuais. O autor aponta que, no Forro, a sílaba é aberta na maioria dos casos, e os casos de sílaba fechada seriam uma influência moderna do Português. Assim, ao fazer uso regular do Forro, o falante estaria mais sujeito à monotongação de /ei/.

Além das variáveis sociais apresentadas, foram controladas ao todo cinco variáveis estruturais, definidas com base nos estudos sobre o tema realizados no PB. Foram elas: *o contexto antecedente*, *o contexto subsequente*, *a tonicidade da sílaba*, *o número de sílabas no vocábulo* e *a posição no vocábulo (radical ou sufixo)*.

Na análise, foram consideradas duas amostras, uma referente ao /ei/ em contexto medial, outra, em contexto final de vocábulo.

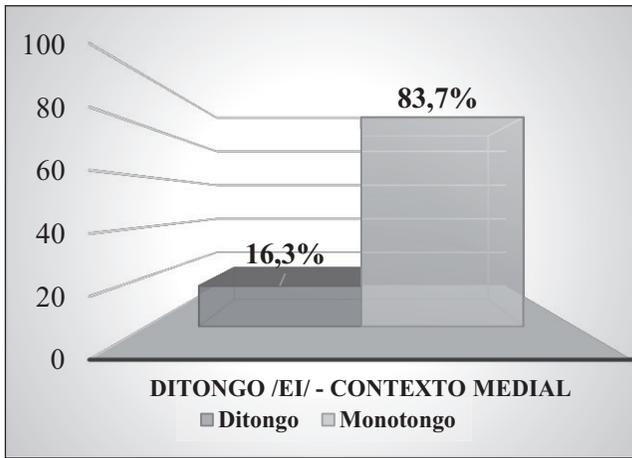
## 5. RESULTADOS DAS ANÁLISES

No levantamento realizado nas 17 entrevistas do *corpus* que serviu de base para o trabalho, foram encontrados 736 dados do ditongo /ei/, tanto em contexto medial de vocábulo, como em “d/ei/xa”, quanto em contexto final em formas verbais, como em *fiq/ei/*. Na amostra geral, registraram-se realizações de /ei/ ditongado – [ej] – ou monotongado – [e] –, não havendo ocorrências de [ɐj], como no PE. Os índices gerais demonstraram que, na variedade urbana do PST, a monotongação de /ei/ é um processo produtivo, com 70,1% de aplicação, correspondentes a 516 das 736 ocorrências, constituindo, portanto, uma regra variável.

Em contexto medial, foram levantadas 417 ocorrências do ditongo, e em contexto final, 319. A seguir, apresentam-se os resultados para cada posição.

### 5.1 O ditongo /ei/ em contexto medial de vocábulo

Em contexto medial, o índice de monotongação se mostrou maior do que o encontrado na análise conjunta, atingindo 83,7% dos dados, conforme exposto na Figura 3 e na Tabela 2.



**Figura 3** Percentuais de monotongação de /ei/ no PST(contexto medial de vocábulo)

**Tabela 2** Ocorrências e percentuais de monotongação de /ei/ no PST (contexto medial de vocábulo)

| REALIZAÇÃO DE /ei/ | OCO     | %    |
|--------------------|---------|------|
| [ej] "dinh[e]ro"   | 68/417  | 16,3 |
| [e] "az[e]te"      | 349/417 | 83,7 |

Na análise, duas variáveis mostraram-se atuantes para a queda da semivogal, sendo uma de caráter social, o nível de *escolaridade*, e uma de natureza estrutural, o *contexto subsequente*. A rodada que propiciou a análise teve significância .000 e *input* .91.

A primeira variável selecionada, escolaridade, demonstrou que há uma maior propensão à monotongação entre os falantes menos escolarizados, fator que apresentou um peso relativo bastante favorecedor, .81. Tal favorecimento decai conforme aumenta o índice de escolaridade dos falantes. Para o grupo intermediário,

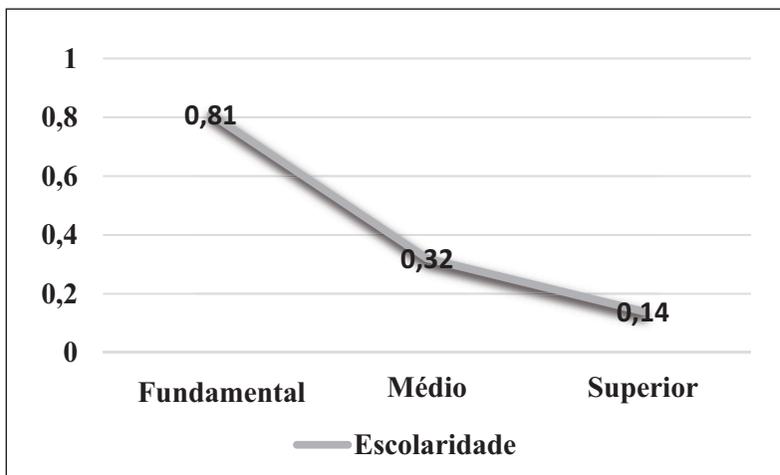
o P. R. foi de .32, já considerado desfavorecedor, e para o grupo de falantes de alta escolaridade o P. R. foi de .14, altamente desfavorecedor do fenômeno.

A Tabela 3 apresenta o número de ocorrências, percentuais e peso relativo de cada fator.

**Tabela 3** Atuação da variável *escolaridade* sobre o processo de monotongação de /ei/ no PST em contexto medial de vocábulo

| FATORES     | OCO     | PERC. | P. R.               |
|-------------|---------|-------|---------------------|
| Fundamental | 179/184 | 97.3% | .81                 |
| Médio       | 111/138 | 80.4% | .32                 |
| Superior    | 59/95   | 62.1% | .14                 |
| Input: .91  |         |       | Significância: .000 |

Tais resultados levam à conclusão de que, quanto maior o contato com a norma escrita, menor a probabilidade de o indivíduo cancelar, na fala, a semivogal. Na Figura 4, é possível observar a curva decrescente do índice de monotongação de acordo com o aumento da escolaridade.



**Figura 4** Atuação da variável nível de escolaridade sobre o processo de monotongação de /ei/ no PST em contexto medial de vocábulo

Este não foi o único trabalho que observou a atuação da variável escolaridade sobre o fenômeno da monotongação de /ei/. Em estudo sobre o PB, Lopes

(2002) observou, além da atuação dos fatores estruturais, como a presença de um tepe e de uma consoante palatal no contexto subsequente, conforme exposto na seção 3.1 deste capítulo, a atuação da variável escolaridade, última variável selecionada em sua análise. Os resultados da autora demonstram um comportamento da escolaridade parecido com os que aqui apresentados, sendo os indivíduos mais escolarizados os que menos monotongam e os menos escolarizados os que mais produzem o fenômeno. Os pesos relativos encontrados por ela foram .66, para a faixa de escolaridade alta, .51, para a média, e .31 para a baixa.

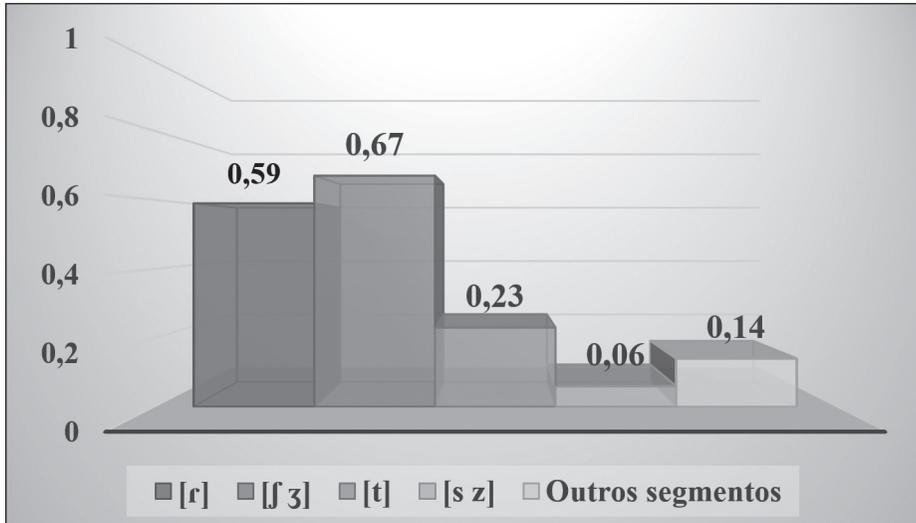
Comportamento similar é também encontrado no trabalho de Silveira (2013) sobre o PST. No estudo do autor, a variável escolaridade foi a penúltima selecionada, e apresentou, da escolaridade mais alta à mais baixa, respectivamente, os pesos relativos .27, .48 e .64. Mais uma vez, quanto maior a escolarização, menor a incidência de queda da semivogal, o que reafirma que, de fato, parece que o contato com a norma escrita exerce influência sobre o fenômeno na fala.

A segunda variável selecionada como atuante sobre a monotongação em contexto medial foi, como dito, o *contexto subsequente*. Tal resultado estava dentro do esperado, devido ao apresentado nas seções anteriores, e os fatores atuantes como favorecedores do processo também foram os previstos: fricativas pós-alveolares e tepe. Seus pesos relativos foram, respectivamente, .67 e .59.

Na Tabela 4, expõem-se os resultados para cada fator da variável e, na Figura 5, destacam-se especificamente os pesos relativos.

**Tabela 4** Atuação da variável contexto subsequente para a monotongação de /ei/ no PST (contexto medial de vocábulo)

| FATORES          | OCO     | PERC  | P. R. | EXEMPLO             |
|------------------|---------|-------|-------|---------------------|
| [r]              | 241/268 | 89.9% | .59   | mad[e]ra            |
| [ʃ ʒ]            | 45/47   | 95.7% | .67   | d[e]xo<br>f[e]jão   |
| [t]              | 56/86   | 65.1% | .23   | f[e]tor (feitor)    |
| [s z]            | 5/12    | 41.7% | .06   | ref[e]ção           |
| Outros segmentos | 2/4     | 50%   | .14   | mant[e]ga           |
| Input: .91       |         |       |       | Significância: .000 |



**Figura 5** Atuação da variável contexto subsequente para a monotongação de /ei/ no PST, expressa só em pesos relativos (contexto medial de vocábulo)

O fator apontado como o maior favorecedor foram as fricativas pós-alveolares, com P. R.: .97 e o elevado percentual de monotongação de 95,7%, o que representa 45 de 47 ocorrências, quase a totalidade. Podemos retomar as observações de Bisol (1989), que diz que diante de consoante palatal a inserção ou apagamento da semivogal não gera consequências para o significado da palavra, pois ela seria um espriamento do traço [+alto] da palatal, gerando o chamado falso ditongo. O resultado deste trabalho parece reafirmar a posição da autora, o que também ocorre com o segundo fator selecionado como favorecedor, comentado a seguir.

O tepe teve peso relativo de .59 e percentual de 89,9%, também bastante elevado, representando 241 ocorrências de um total de 268. Bisol aponta que a monotongação nesse contexto igualmente não gera danos ao significado da palavra, sendo também um caso de falso ditongo, não presente na estrutura subjacente da sílaba, o que explica a elevada aplicação do fenômeno de queda da semivogal.

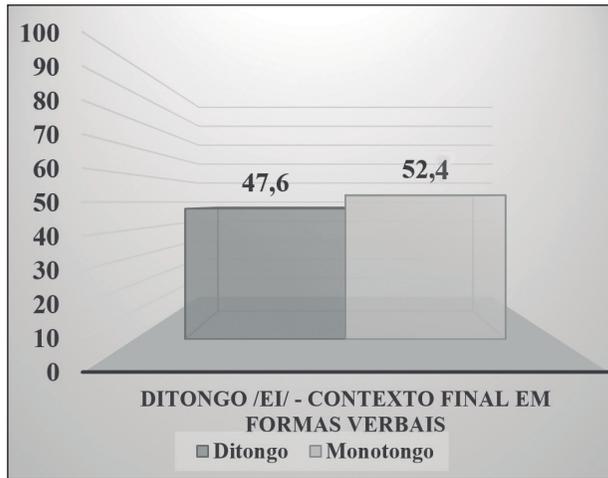
Quase todos os demais contextos funcionaram de maneira similar ao que ocorre no PB, ou seja, são desfavorecedores do processo. Eles foram divididos em *fricativas alveolares* e *outros segmentos*, que compreendeu apenas 4 ocorrências, sendo um de [n] (“treinarem”), uma de [f] (“gaseificados”), uma de [d] (“Almeida”) e uma de [g] (“leigo”). Os pesos relativos desses fatores foram, respectivamente, .06 e .14. A exceção para essa semelhança foi a presente de um [t] no contexto subsequente.

Embora, a princípio, se observado apenas o P. R., que foi de .23, tal contexto se mostre desfavorecedor, se se observa o percentual de ocorrência do processo,

que foi de 65,1%, representando 56 de 86 dados, verifica-se que a monotongação de /ei/ pode ocorrer diante de [t] no PST, o que não ocorre no PB, indicando uma divergência entre as duas variedades.

## 5.2 O ditongo /ei/ em contexto final

Os resultados para a análise do ditongo /ei/ em contexto final em formas verbais apresentaram o índice percentual mais baixo em comparação ao índice geral e ao do contexto medial, tendo a monotongação ocorrido em 52, 5% dos dados, como demonstram o gráfico e a tabela a seguir.



**Figura 6** Percentuais de monotongação de /ei/ no PST (contexto final em formas verbais)

**Tabela 5** Ocorrências e percentuais de monotongação de /ei/ no PST (contexto final em formas verbais)

| REALIZAÇÃO DE /ei/ | OCO     | %    |
|--------------------|---------|------|
| [ej] "deixei"      | 152/319 | 47,6 |
| [e] "fiquei"       | 167/319 | 52,4 |

Ainda que o percentual seja mais equilibrado, é possível afirmar que a monotongação em contexto final em formas verbais ocorre no PST, diferentemente do que se observa do PB. Paiva (1996), como já se salientou, não chega a considerar o contexto final em seu estudo pois considera que a manutenção da semivogal nesse contexto seria categórica, fato que pode ser observado também no trabalho de Ribeiro (1990), que, ao analisar a variável *tipo de sílaba*, dividida entre os fa-

tores *não final* e *final*, encontrou como resultado 0% de queda da semivogal nesta última. Assim, este caso constitui um traço divergente do PST em relação ao PB.

A única variável apontada pelo Programa Goldvarb-X como atuante para o processo foi *frequência de uso de um crioulo*. Os resultados demonstraram que o fator *frequência alta* seria o grande favorecedor da queda da semivogal, com P. R.: .93. Já para as frequências *média* e *baixa* os pesos relativos estão próximos da neutralidade, ainda que esta última tenha um índice um pouco mais inibidor, .48 e .43, respectivamente. O total de ocorrências e percentuais de monotongação por fator são apresentados na Tabela 6.

**Tabela 6** Atuação da variável *frequência de uso de um crioulo* para a monotongação de /ei/ no PST (contexto final em formas verbais)

| FATORES     | OCO          | PERC.               | P. R.      |
|-------------|--------------|---------------------|------------|
| Baixa       | 78/164       | 47.6%               | .43        |
| Média       | 71/136       | 52.2%               | .48        |
| <b>Alta</b> | <b>18/19</b> | <b>94.7%</b>        | <b>.93</b> |
| Input: .54  |              | Significância: .000 |            |

Cabe lembrar que a variável foi formulada com base na hipótese de que quanto maior uso um indivíduo fizesse de um crioulo, mais sua fala estaria sujeita a influências dessa língua e, conseqüentemente, mais se afastaria da norma do PE. Ferraz (1979) apresenta um estudo sobre o Forro, no qual, além de tratar da sua formação, apresenta as suas principais características fonológicas e gramaticais.

O autor demonstra que o principal padrão silábico do Forro é do tipo CV, seguido dos padrões V e CVC, sendo este menos usual que o anterior. Segundo ele, a estrutura da sílaba do Forro apresenta uma simplificação com relação à do PE, pois no Forro a sílaba é predominantemente aberta, sendo os casos de sílaba fechada, no padrão CVC, uma influência do Português.

O autor afirma, no entanto, que, no Forro, o *onset* da sílaba pode ser constituído por uma seqüência de até três segmentos consonantais. No caso de uma seqüência de duas consoantes, a segunda equivaleria a /l/ ou a uma semivogal, havendo também a possibilidade de a primeira consoante ser uma homorgânica nasal, como no exemplo apresentado pelo autor “*nda*”(andar). Já os casos em que o *onset* apresenta três consoantes seriam bem específicos: “(...) C1 seria sempre igual a uma fricativa sibilante surda, C2 a [t] ou [k], e C3 a /l/” (1979, p. 27).

Ainda em relação ao resultado da análise aqui apresentada, uma observação se faz necessária. Apenas um dos informantes da pesquisa declarou fazer alto uso do Forro. Portanto, todos os dados referentes ao fator *alta frequência*, que apresentou um total de 19 ocorrências, das quais 18 com queda da semivogal, pertencem ao mesmo indivíduo, da faixa etária mais alta e de menor nível de escolaridade, fator este que, em contexto medial, se mostrou favorecedor da monotongação.

Apesar do exposto, o resultado não contraria a hipótese formulada, ao contrário, mostra evidências de que, em análises futuras, com um corpus ampliado e um maior número de informantes com o mesmo perfil, resultado semelhante poderá ser encontrado.

Por fim, ainda que essa variável tenha sido a única selecionada pelo programa, vale a pena mencionar o resultado referente à variável *nível de escolaridade*. Em termos percentuais, observa-se a mesma curva decrescente obtida na análise do contexto medial, sendo os indivíduos menos escolarizados os que mais aplicam a monotongação, seguidos pelos da faixa intermediária e pelos da faixa alta, com percentuais de 60,6%, 50% e 46,2%, respectivamente. Isso parece indicar que realmente há a influência dessa variável na fala da comunidade estudada.

## 6. CONSIDERAÇÕES FINAIS

Neste estudo apresentaram-se os resultados da dissertação de Passos (2018) sobre a realização do ditongo /ei/ na variedade urbana do PST, na qual se demonstra que a monotongação é um processo produtivo e constitui uma regra variável na comunidade.

As restrições que favorecem a aplicação dessa regra variam segundo a posição da sílaba no vocábulo. Em posição medial, é condicionada pelo nível de escolaridade e o contexto subsequente, sendo os indivíduos menos escolarizados os que mais a implementam, sobretudo quando o ditongo é seguido por fricativa palato alveolar ou por tepe. Em contexto final, verificou-se que indivíduos que se utilizam do Forro, com alta frequência, em suas práticas intercomunicativas são os mais suscetíveis à monotongação.

Os resultados demonstraram haver convergências entre as variedades urbanas do PST e do PB quanto ao contexto medial, apesar de no PB a monotongação não ocorrer diante de [t]. Já com relação ao contexto final, observa-se divergência, uma vez que a monotongação nessa posição não é típica no PB.

Por fim, espera-se que este estudo possa contribuir para o conhecimento dessa variedade africana do Português, conhecendo-a um pouco mais do ponto de vista fonético-fonológico, e, assim, contribuir também para o conhecimento do PB.

## 7. REFERÊNCIAS

- AVELAR, Juanito; GALVES, Charlotte. O papel das línguas africanas na emergência da gramática do português brasileiro. *Linguística*, 30 (2), diciembre 2014: 241-288.
- BISOL, Leda. O ditongo na perspectiva da fonologia atual. *Revista DELTA*, São Paulo, v. 05, n. 2, p. 185-224, 1989.
- BISOL, Leda. Ditongos derivados. *Revista DELTA*, v. 10, n. especial, p. 123-140, 1994.
- BISOL, Leda. Ditongos derivados: um adendo. In: LEE, S. H. (Org.). *Vogais além de Belo Horizonte*. Belo Horizonte: FALE/UFMG, 2012.
- BRANDÃO, Sílvia Figueiredo. Concordância nominal na variedade não standard do português falado em São Tomé. Comunicação apresentada ao XVI Congresso Internacional de ALFAL, Alcalá de Henares, Espanha. Universidad de Alcalá, 6-9 jun., 2011.
- BRANDÃO, Sílvia Figueiredo. Patterns of plural agreement within the Noun Phrase. *Journal of Portuguese Linguistics*, 12 (2): 51-100, 2013.
- BRANDÃO, Sílvia Figueiredo; VIEIRA, Sílvia Rodrigues. A concordância nominal e verbal no Português do Brasil e no Português de São Tomé: uma abordagem sociolinguística. *Papia: Revista Brasileira de Estudos Crioulos e Similares*, v. 22, n. 1, p. 7-39, 2012.
- CALLOU, Dinah; LEITE, Yonne. *Iniciação à fonética e à fonologia*. 11. ed. Rio de Janeiro: Zahar, 2009. [1990].
- CÂMARA JR., Joaquim Mattoso. *Estrutura da Língua Portuguesa*. 44. ed. Rio de Janeiro: Vozes, 2011. [1970].
- CINTRA, Lindley. *Estudos de dialectologia portuguesa*. Lisboa: Sá da Costa, 1983.
- CUNHA, Celso. Conservação e inovação no Português do Brasil. *O Eixo e a Roda*. Belo Horizonte, 5: 199-230, 1986.
- FERRAZ, Luiz Ivens. *The creole of São Tomé*. Johannesburg: Witwatersrand University Press, 1979.
- GONÇALVES, C. A. V. Ditongos decrescentes: variação & ensino. *Revista de Estudos da Linguagem*. Belo Horizonte: Faculdade de Letras da UFMG, v. 6. n. 5. p. 159-192, jan./jun. 1997.
- GONÇALVES, C. A. V.; COSTA, R. G. R. Sobre a interpretação fonológica dos ditongos em português. In: *IV Congresso da Assel-Rio*, 1995, Niterói. Anais do IV Congresso da Assel-Rio. Niterói: Assel-Rio, 1994. v. 1. p. 76-85.
- LABOV, William. *Padrões sociolinguísticos*. Trad.: Marcos Bagno; Marta Scherre e Caroline Cardoso. São Paulo: Parábola, 2008. [1972].
- LABOV, W. *Principles of linguistic change: internal factors*. Oxford: Blackwell Publishers, 1994. v. 1.
- LABOV, W. *Principles of linguistic change: social factors*. Oxford: Blackwell Publishers, 2001. v. 2.

LOPES, Raquel. A realização variável dos ditongos /ow/ e /ej/ no português falado em Altamira/PA. 97 f. Dissertação (Mestrado em Linguística), UFPA, Belém, PA, 2002.

MATEUS, Maria Helena Mira; D'ANDRADE, Ernesto. The phonology of portuguese. Oxford: University Press, 2000.

PAIVA, M. C. A. Supressão das semivogais nos ditongos decrescentes. In: SILVA, G. M. O.; SCHERRE, M. M. P. (Org.). *Padrões sociolinguísticos*. Rio de Janeiro: Tempo Brasileiro, 1996a. p. 217-236.

PAIVA, M. C. A. O percurso da monotongação de [ey]: observações no tempo real. In: PAIVA, M. C.; DUARTE, M. E. L. (Org.). *Mudança linguística em tempo real*. Rio de Janeiro: Contra Capa, 2011. p. 31-46.

PASSOS, R. R.; BARCELOS, N. H. M. O ditongo /ei/ na fala do Português de São Tomé. Comunicação apresentada à XXXVII Jornada de Iniciação Científica da UFRJ. Faculdade de Letras, 2015.

PASSOS, R. R. *O Ditongo /ei/ na variedade urbana do Português de São Tomé*. 2018. 88 fls. Dissertação de Mestrado em Letras Vernáculas, Língua Portuguesa. Faculdade de Letras, Universidade Federal do Rio de Janeiro, 2018.

PEREIRA, Gerusa. *Monotongação dos ditongos /aj/, /ej/, /ow/ no Português falado em Tubarão (SC): estudo de casos*. Dissertação (Mestrado em Ciências da Linguagem) – Tubarão-SC, Universidade do Sul de Santa Catarina, 2004. 134f.

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

RIBEIRO, Denise Aparecida Sofiati de Barros. *O apagamento dos ditongos decrescentes orais no sudoeste do Paraná*. Dissertação (Mestrado em Linguística), UFPR, Curitiba, 1990. 125f.

SILVA, Vanessa Meireles de Oliveira. *Analyse phonologique et métrique des glides et diphtongues en portugais brésilien*. Thèse de Doctorat. Paris: Université Paris 8 Vincennes – Saint-Denis, 2014.

SILVEIRA, A. C. *Ditongos no Português de São Tomé e Príncipe*. 2013. 180fls. Dissertação (Mestrado em Filologia e Língua Portuguesa). Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, 2013.

TOLEDO, Eduardo Elisal de. 2011. 109f. *A monotongação do ditongo decrescente [ej] em amostra de recontato de Porto Alegre*. Dissertação (Mestrado em Linguística), PUCRS, Porto Alegre, RS, 2011.

VIEIRA, S. R.; BAZENGA, A. M. Patterns of 3rd person verbal agreement. *Journal of Portuguese Linguistics*, v. 12, p. 7-50, 2013.

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



# VARIÁVEIS MORFOSSINTÁTICAS



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

SILVIA FIGUEIREDO BRANDÃO

## 1. INTRODUÇÃO: POR QUE FOCALIZAR VARIEDADES AFRICANAS DO PORTUGUÊS

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

### (a) *São Tomé e Príncipe*

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

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

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

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

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

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

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

Sobre esse procedimento Figueiredo conclui que:

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

*(b) Moçambique*

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

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

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

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

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

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

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

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

Quanto às duas variáveis sociais mais relevantes, ela conclui que

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

(c) *Cabo Verde*

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

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

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

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

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

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

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

(d) *Angola*

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

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

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

### 3. OBJETIVOS E ORIENTAÇÃO TEÓRICO-METODOLÓGICA

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

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

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

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

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

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

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

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

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

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

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

## 5. A VARIEDADE URBANA DO PORTUGUÊS DE SÃO TOMÉ

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

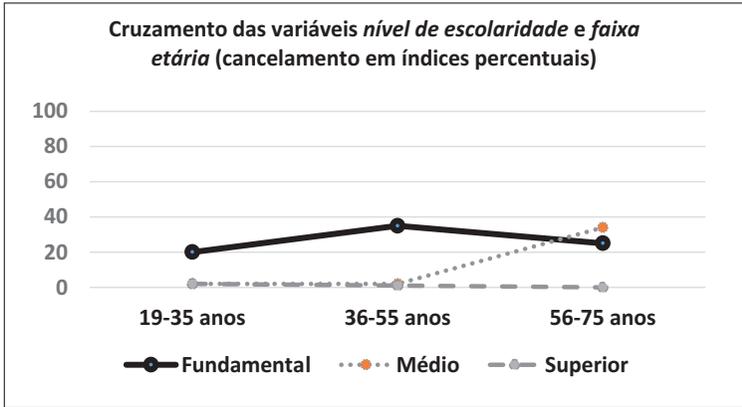
Na Amostra 1, obtiveram-se 2439 ocorrências de aplicação da marca (93,4%) e apenas 173 (6,6%) de cancelamento. Os resultados indicaram o *nível de escolaridade* e a *posição linear e relativa dos constituintes no SN* como as duas variáveis mais relevantes para o cancelamento ou não da marca de número,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

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

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

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

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



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

Fonte: Brandão, 2011b

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Fonte: Brandão, 2011c

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

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

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

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

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

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

Fonte: Brandão (2011c)

Independentemente da distribuição assimétrica, verifica-se (Tabela 5) que são os indivíduos que se comunicam preferencialmente num crioulo os que mais tendem a não implementar a marca de número (P. R.: .75). Os três têm nível fundamental de instrução: duas mulheres, uma da faixa mais jovem, outra da mais velha e um homem, também desta última faixa etária. Com esses indivíduos, contrastam os que nunca ou pouco se expressam em crioulo (P. R.: .41) e os que o utilizam apenas eventualmente (P. R.: .52). Observe-se, ainda, que o cancelamento também incide de forma escalar a depender da frequência de uso do crioulo: alta > .75; média > .52; baixa/zero > .41.

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

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

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

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

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

Fonte: Brandão (2011c)

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

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

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

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

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

Fonte: Brandão, 2011c

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

## 6. A VARIEDADE URBANA DO PORTUGUÊS DE MOÇAMBIQUE

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

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

**Quadro 2** Distribuição dos informantes pelas variáveis sociais

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

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

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

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

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

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

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

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

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

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

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

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

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

## 6.1 A análise geral quantitativa

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

## 6.2 Estendendo a análise: outros aspectos quantitativo-qualitativos

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

(12) [das *família*] principalmente

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

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

(14) só assistindo [nos *televisor*]

(15) nós subimos [aqueles *carro*]

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

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

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

(19) [os *livro*] dão na escola

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

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

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

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

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

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

(26) ir à igreja n[os *domingo*] ....rezar

(27) [dezoito *hora*] estão em casa

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

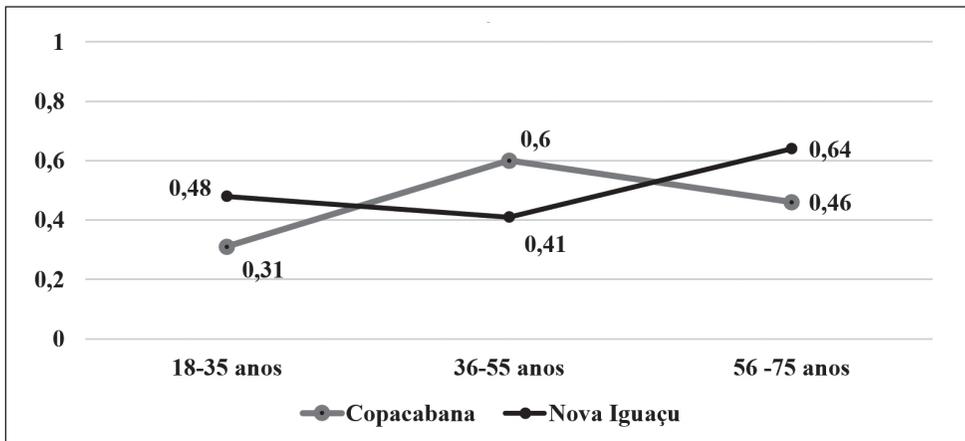
| PB                              |             |      |       |            |      |       |
|---------------------------------|-------------|------|-------|------------|------|-------|
| NÍVEL DE ESCOLARIDADE           | NOVA IGUAÇU |      |       | COPACABANA |      |       |
|                    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

Fonte: Brandão (2013: 72), com alterações

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

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

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



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

Fonte: Brandão (2013: 82), com alterações

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

## 8. SÍNTESE DOS RESULTADOS: ALGUMAS REFLEXÕES

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

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

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

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

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

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

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

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

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

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

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

[+marcas] → → [-marcas]

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

Fonte: Brandão (2013), com acréscimos

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

[...] a primeira posição linear no pré-núcleo tende a ser mais marcada, o que sugere seja esse o *locus* por excelência da marca. A partir do núcleo em segunda posição, vai decrescendo gradativamente a sua presença. [...] No esquema proposto, a linha vertical contínua que separa o pré-núcleo do núcleo representaria o limite do espaço mais saliente para a marcação (o *locus* esquerdo), enquanto a linha vertical pontilhada entre o núcleo e o pós-núcleo indicaria que essas áreas não sofreriam, quanto à tendência à não marcação, solução de continuidade. Observe-se, como se ressaltou antes, que o deslocamento para a primeira posição de um núcleo ou de um constituinte que ocupe posição pós-nuclear concorrerá para que se torne mais suscetível a receber marca de número.

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

É grande a complexidade também em São Tomé e Príncipe, em que o Português coexiste com crioulos de base lexical portuguesa, num espaço geográfico exíguo se comparado ao de Moçambique. No primeiro dos citados países, num total de 1001km<sup>2</sup>, vivem 178.739 indivíduos, enquanto, no segundo, 28.861.863 habitantes espalham-se por uma área de 801.590km<sup>2</sup>, 1.257.453 deles na capital Maputo.

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

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

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

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

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

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

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

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

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

Fonte: Brandão (2013: 90), com alterações

## 9. CONSIDERAÇÕES FINAIS

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

Como observa Galves (2008: 161):

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

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

## REFERÊNCIAS

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

- AMARAL, Amaral. *O dialeto caipira*. 3. ed. São Paulo: Hucitec, Secretaria de Cultura, Ciência e Tecnologia, 1976 [1920].
- AVELAR, Juanito; GALVES, Charlotte. O papel das línguas africanas na emergência da gramática do português brasileiro. *Linguística*, 30 (2): 241-288, 2014.
- BAXTER, Alan. The development of variable NP plural agreement in a re-structured African variety of Portuguese. In: ESCURE, G.; SCHWEGLE, A. *Creoles, contact, and language change*. Amsterdam/Philadelphia: John Benjamins, 2004. p. 97-126.
- BAXTER, Alan. A concordância de número. In: LUCCHESI, D.; BAXTER, A.; RIBEIRO, I. (Org.). *O português afro-brasileiro*. Salvador: EDUFBA, 2009. p. 269-293.
- BRAGA, Maria Luisa. *A concordância de número no sintagma nominal no Triângulo Mineiro*. 88f. Dissertação (Mestrado em Língua Portuguesa) – Pontifícia Universidade Católica, Rio de Janeiro, 1977.
- BRAGA, Maria Luisa; SCHERRE, Maria Marta Pereira. A concordância de número no SN na área urbana do Rio de Janeiro. In: ENCONTRO NACIONAL DE LINGÜÍSTICA, 1, 1976, Rio de Janeiro. *Anais ...* Rio de Janeiro: PUC, 1976. p. 464-477.
- BRANDÃO, Sílvia Figueiredo. Concordância nominal em duas variedades do português: convergências e divergências. *Revista Veredas*, 15 (1): 164-178, 2011a.
- BRANDÃO, Sílvia Figueiredo. Concordância nominal na variedade urbana do Português falado em São Tomé. *Ciências Sociais e Humanas em Revista*, 33(1): 31-42, 2011b.
- BRANDÃO, Sílvia Figueiredo. O cancelamento da marca de número nominal na variedade urbana não *standard* do Português de São Tomé. Comunicação apresentada ao XVI Congreso Internacional de la Asociación de Lingüística y Filología de la América Latina (ALFAL). Alcalá de Henares, Madrid, 6-9 de junho, 2011c.
- BRANDÃO, Sílvia Figueiredo. Patterns of agreement within the Noun Phrase. *Journal of Portuguese Linguistics*, 12(2): 51-100, 2013.
- BRANDÃO, Sílvia Figueiredo; VIEIRA, Sílvia Rodrigues. A concordância nominal e verbal no Português do Brasil e no Português de São Tomé: uma abordagem sociolinguística. *Papia: Revista Brasileira de Estudos Crioulos e Similares*, 22(2): 7-41, 2012a.
- BRANDÃO, Sílvia Figueiredo; VIEIRA, Sílvia Rodrigues. Concordância nominal e verbal: contribuições para o debate sobre o estatuto da variação em três variedades urbanas do português. *Alfa, Revista de Linguística*, 56(3): 1035-1064, 2012b.
- BRANDÃO, Sílvia Figueiredo; VIEIRA, Sílvia Rodrigues. Para uma análise qualitativa de dados quantitativos: o caso da concordância em Português. Comunicação apresentada ao IV DIA: Repenser la variation linguistique. Zurique: Universidade de Zurique, 12-14 de setembro, 2016.
- FIGUEIREDO, Carlos. Filipe Guimarães. A concordância variável no sintagma nominal plural do português reestruturado de Almoxarife. *Papia: Revista Brasileira de Estudos Crioulos e Similares* 18: 23-44, 2008.
- FIGUEIREDO, Carlos. Filipe Guimarães. 2 v. 798 fls. *A concordância plural variável no sintagma nominal do português reestruturado da comunidade de Almoxarife, São Tomé*

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

# PADRÕES DE CONCORDÂNCIA VERBAL DE TERCEIRA PESSOA PLURAL NO PORTUGUÊS DE SÃO TOMÉ E NO PORTUGUÊS DE MOÇAMBIQUE

**KAREN CRISTINA DA SILVA PISSURNO**

## **1. INTRODUÇÃO**

Os estudos sociolinguísticos sobre a concordância verbal, iniciados por volta de 1970 nas amostras brasileiras, demonstram uma variação muito produtiva em relação ao uso das regras de marcação plural no Português do Brasil (PB), especialmente por conta de fatores extralinguísticos, como os relativos a elementos socioeconômicos, à escolaridade dos indivíduos ou, ainda, às características de suas comunidades de fala, sobretudo em relação a seu perfil, mais rural ou mais urbano. No que se refere às variedades africanas, entretanto, não se dispõe de descrições científicas suficientes, sobretudo ancoradas em abordagens quantitativas de dados, que constatem o estatuto da regra de concordância.

Acredita-se, portanto, na relevância atribuída às investigações sociolinguísticas das variedades de Português do continente africano que considerem a marcação de número, não somente por contribuírem para a descrição das diferentes normas da Língua Portuguesa, mas, especialmente, por abordarem a questão da intensa situação de contato linguístico presente nessas sociedades. Tais investigações podem colocar em relevo especificidades que são encontradas em contexto multilíngue e que tendem, invariavelmente, a distanciar, de certa maneira, as

variedades africanas do seu suposto “modelo” de aquisição – o Português Europeu (PE) –, segundo estudos que revelam a alta preferência dessa variedade pelo uso das marcas de plural, independentemente de estrato social ou escolaridade.

Além disso, o aprofundamento das questões relacionadas ao contato linguístico pode trazer luzes, em última instância, ao produtivo debate quanto às interpretações sobre as origens e as especificidades do PB. De acordo com Lucchesi; Baxter; Silva (2009), teria ocorrido na formação do Português em terras brasileiras o que se pode definir como um caso de transmissão linguística irregular, segundo o qual os modelos linguísticos adquiridos pelos novos falantes de Português indicariam o uso de adaptações representativas do contato entre línguas, dentre as quais é possível destacar a simplificação morfológica – o que se correlacionaria à marcação de pluralidade. Tendo em vista o cenário descrito pelos autores, entende-se que as características de formação das variedades africanas em meio ao multilinguismo – que, guardadas suas proporções, se assemelham à formação do PB – possam colaborar, considerando o Princípio do Uniformitarismo, para a explicação dos processos que originaram a variedade brasileira do Português.

Sendo assim, pretende-se considerar resultados de investigações sobre a concordância verbal em duas variedades africanas do Português, a de São Tomé e Príncipe (PST) e a de Moçambique (PM), a partir dos princípios da Teoria da Variação e Mudança (WEINREICH, LABOV, HERZOG, 1968; LABOV, 1972, 1994, 2003) e verificar se, nessas variedades, a concordância verbal constitui uma regra semicategórica ou variável, nos termos de Labov (2003). Assim, a partir da distribuição geral dos dados quanto ao fenômeno em questão e da investigação das motivações linguísticas e extralinguísticas que (des-)favorecem a realização da marca de terceira pessoa do plural nas formas verbais, constata-se o tipo de regra presente em cada variedade do Português. Sobre os fatores extralinguísticos, considera-se de extrema importância, devido ao contexto multilíngue das regiões examinadas, observar a influência do uso simultâneo de outra(s) língua(s) dominada(s) pelos falantes de Português sobre os resultados obtidos.

Para tanto, o presente capítulo sintetiza, na Seção 2, as informações a que se teve acesso quanto aos padrões de concordância nas referidas variedades africanas e suas diferenças em relação ao PE. Na Seção 3, apresentam-se os objetivos de duas pesquisas da concordância verbal – uma com dados do PM (PISSURNO, 2017) e a outra, do PST (VIEIRA; SILVA, 2017) – no âmbito do Projeto *Estudo comparado dos padrões de concordância em variedades africanas, brasileiras e europeias do Português*, além dos procedimentos metodológicos utilizados para a sua consecução. Em seguida, nas Seções 4 e 5, os resultados dos referidos estudos são apresentados, para que, na Seção 6, seja feita uma breve comparação entre os índices alcançados. Por fim, são tecidas, na Seção 7, as considerações finais.

## 2. O PE E A CONCORDÂNCIA VERBAL NAS VARIEDADES AFRICANAS DO PORTUGUÊS

Tendo em vista o processo de colonização em terras africanas, o PE tem atuado como norma de referência e modelo de aquisição no contexto de formação das variedades africanas da Língua Portuguesa, conforme propõe Pissurno (2017):

O Português Europeu é a variedade tida como modelo de formação para todas as outras variedades do Português, por ter-se configurado historicamente como língua de maior prestígio, representante do colonizador e das oportunidades de ascensão social. Por representar a idealização do que seria o padrão correto a ser seguido, o PE é, muitas vezes, considerado como uma língua em que a variação é inexpressiva, especialmente no que tange ao fenómeno da concordância verbal, porque sua ocorrência seria ínfima. (PIS-SURNO, 2017, p. 88)

Sendo assim, torna-se imprescindível o conhecimento dos padrões de concordância na variedade europeia do Português para a compreensão das semelhanças e diferenças entre ela e as novas variedades em formação ou já configuradas. Investigações variacionistas com dados do PE (VAREJÃO, 2006; MONGUI-LHOTT, 2009; GANDRA, 2009; VIEIRA, 2011; RUBIO, 2012; MONTE, 2012; VIEIRA; BAZENGA, 2013) constataram que, efetivamente, ocorre ampla preferência pela marcação de pluralidade no PE. Considerando a aplicação das regras de concordância no PE, verifica-se que os índices de marcação em verbos de 3ª pessoa plural variam entre 90% e 97% das ocorrências computadas – variação que, segundo Vieira; Bazenga (2013), é motivada pelas opções metodológicas de cada estudo quanto aos critérios relativos à coleta de dados. Na realidade, trata-se de uma regra semicategórica de concordância, ficando as ocorrências de não marcação restritas a contextos muito específicos, a saber: sujeito posposto ou não realizado, de referência inanimada ou, ainda, representado por pronome relativo *que*. Em termos extralinguísticos, destaca-se, ainda, que a variável escolaridade não impõe qualquer restrição sobre os resultados europeus, ao contrário do que acontece, ao que tudo indica, nas demais variedades.

Do perfil da variedade europeia, deriva que os padrões de concordância verificados figurem como modelo de aprendizagem do Português, a ser apresentado nos manuais (gramáticas e livros didáticos) das escolas africanas, de modo que todo tipo de desvio a esse padrão seja possivelmente considerado como estrutura a ser corrigida.

Em relação ao PM, não se teve acesso a trabalhos variacionistas sobre a marcação de plural em formas verbais com base em análises contrastivas do Portu-

guês. Gonçalves (2015), tratando especificamente desse fenômeno morfossintático, verificou que até mesmo os trabalhos que se dedicaram à descrição de dados parecem ter uma preocupação sistemática em indicar qual é a maneira correta de realizar tais estruturas. Como afirma a autora:

A maior parte dos estudos sobre a CV em PM, num total de 8, foram realizados numa perspectiva descritiva e, em alguns casos, têm como horizonte intervenções didáticas, destinadas a promover a competência gramatical dos estudantes nesta área gramatical. Esta dimensão didática decorre do fato de que, em Moçambique, tal como acontece no Brasil, a não realização da regra de concordância verbal constitui “um traço de diferenciação social, de cunho estigmatizante” (Vieira, 2007, p. 85). Até ao momento presente, não foram realizados estudos quantitativos com o objetivo de descrever detalhadamente o padrão da concordância variável no SV. (GONÇALVES, 2015, p. 13)

Gonçalves (1997), no trabalho intitulado “Tipologia de erros do Português Oral de Maputo”, teve por objetivo reunir dados da fala moçambicana que servissem como fonte para a descrição dessa variedade. Da observação desses dados, a autora relata alguns casos de ausência de marcas no verbo, especialmente em contextos de sujeito posposto. Por sua vez, António (2011), considerando textos escritos por universitários, apresenta 42 ocorrências sem marcas, principalmente quando os sujeitos são complexos, nulos ou pronomes relativos, para além da posposição do sujeito – o que sugere um quadro variável de concordância verbal no PM.

Nhongo (2005), Bavo (2011) e António (2011), além de apresentarem, basicamente, os mesmos contextos que mais favorecem a não marcação supracitados, listam as estratégias de ensino e aprendizagem de Língua Portuguesa em Moçambique, fornecendo, inclusive, exercícios a serem trabalhados para o melhor treinamento das regras de concordância em sala de aula. Já Justino (2015a) reforça os condicionamentos para a ausência de marcas encontrados em trabalhos anteriores e indica que o grau de escolaridade dos falantes moçambicanos, nesses contextos, não altera efetivamente os resultados quanto à concordância verbal, pois até mesmo os colaboradores com ensino superior mostram as mesmas tendências de não marcação de plural nos contextos de frases complexas e sujeitos pospostos.

Como se pode observar, a concordância verbal na variedade moçambicana é, de fato, uma área que gera conflitos em relação ao Português Europeu, já que os trabalhos supracitados apontam a existência de contextos que distanciam, em certa medida, o PM de seu suposto modelo. Essa constatação é confirmada no primeiro trabalho sociolinguístico quantitativo sobre a concordância verbal no

Português moçambicano, realizado por Pissurno (2017), cujos resultados serão discutidos no âmbito do presente artigo (cf. Seção 4).

Considerando a variedade são-tomense do Português, algumas investigações sobre a concordância verbal já realizadas (FIGUEIREDO, 2009, 2010; VIEIRA, 2011; BRANDÃO; VIEIRA, 2012, VIEIRA; BAZENGA, 2013, VIEIRA; BRANDÃO, 2014) revelam tendências de que, embora haja ampla preferência pela marcação de plural, a aplicação das regras de concordância na variedade são-tomense pode ser considerada parcial. Observa-se, especialmente, que, resguardadas as diferenças entre os níveis de variação e os contextos em que ela ocorre, São Tomé parece assumir um comportamento de regra variável em contextos semelhantes aos que se verificam nas variedades brasileiras.

Em termos comparativos, Vieira; Brandão (2012b) propõem que os padrões de concordância verbal do PST permitem localizá-lo quanto a seu perfil entre o PB e o PE. Se, de um lado, o PST não alcança níveis de marcação tão altos ou semicategoricos quanto os da variedade europeia, não apresenta, de outro, uma gama tão variada de contextos desfavorecedores da marcação de número nem níveis tão baixos de variação quanto os do PB, sobretudo em variedades urbanas populares.

### 3. OBJETIVOS E PROCEDIMENTOS METODOLÓGICOS

Os estudos das variedades africanas do Português apresentados no presente texto tomam por base os princípios da Teoria da variação e mudança (WEINREICH, LABOV, HERZOG, 1968). Seguindo tais princípios, as investigações descrevem – com base no instrumental estatístico fornecido pelo pacote de programas Goldvarb-X para a realização da análise multivariada dos dados – os fatores linguísticos e sociais relevantes ao condicionamento da marcação de plural em amostras de cada variedade.

Assim, tanto em referência a Moçambique quanto a São Tomé, os trabalhos tratam de todas as ocorrências com e sem marca verbal de terceira pessoa do plural e das restrições para o uso de cada variante. Os resultados relativos a Moçambique foram retirados da pesquisa de mestrado de Pissurno (2017), que considera 18 entrevistas sociolinguísticas realizadas no ano de 2016 na capital do país, Maputo. Já para São Tomé, relatam-se os resultados apresentados por Vieira; Silva (2017), em trabalho que reúne análises anteriores (cf., dentre outros, VIEIRA, 2011; VIEIRA; BRANDÃO, 2014; VIEIRA; BAZENGA, 2013, 2015) considerando 27<sup>1</sup> inquéritos da variedade urbana do PST.

---

<sup>1</sup> É importante salientar que, no caso de São Tomé, não foi possível considerar uma informante idosa de curso superior, porque a amostra VAPOR não apresenta qualquer gravação com esse

Dessa forma, para as duas amostras em estudo, os indivíduos deveriam preencher os critérios estabelecidos em relação à idade (18-35 anos; 36-55 anos; acima de 56 anos), à escolaridade (ensino fundamental; ensino médio; ensino superior) e ao sexo (feminino; masculino).

Ressalte-se que, enquanto na amostra são-tomense todos os informantes se declaram falantes de Português como L1, embora não o empreguem com a mesma frequência nem na mesma variedade de interações sociais, na amostra moçambicana tal configuração não pôde ser mantida, devido ao caráter social bastante distinto constatado na comunidade em questão. Sendo assim, todos os informantes do PM são falantes de Língua Portuguesa, seja ela sua língua materna ou não. A esse respeito, cabe observar que Maputo, sendo a capital do país, recebe pessoas de diferentes origens, advindas de todas as províncias de Moçambique. Esse complexo quadro linguístico dificultou a seleção de informantes que tivessem somente o Português como sua língua materna, como explica Pissurno (2017):

Por conta disso, apesar de buscarmos inquirir pessoas nascidas e criadas na cidade de Maputo, muitas entrevistas, na verdade, contemplam indivíduos nascidos em outras províncias e/ou que têm pelo menos um membro de sua família, seja o pai, a mãe ou o cônjuge, que advém de alguma delas. Tal situação afeta diretamente a questão das línguas faladas por esses indivíduos, já que em cada província existe uma língua local que é mais utilizada, para além do Português, língua oficial do país. Em outras palavras, cada indivíduo que passa a morar na capital carrega consigo as influências de sua própria língua, além de adquirir, independentemente do nível de proficiência, ao menos uma das principais línguas faladas em Maputo, Changana e Rhonga<sup>2</sup>. (PISSURNO, 2017, p. 101).

Além das variáveis referentes à constituição da amostra (idade, escolaridade, sexo), investigou-se, no caso das duas variedades africanas em questão, a influência do uso de outra(s) língua(s) pelo informante, para além do uso de Português, à luz do grupo de fatores idealizado por Brandão (2011). Esse grupo de fatores tem como

---

perfil, motivo pelo qual se chega ao total ímpar de informantes da amostra básica (17 ao invés de 18). Além dessas 17 entrevistas, o estudo de Vieira; Silva (2017) acrescenta os dados de uma amostra complementar, que possui 10 informantes jovens (5 meninos e 5 meninas, entre 10 e 19 anos) ainda em fase escolar, totalizando 27 entrevistados. Ao que tudo indica, esse acréscimo pode permitir a observação do fenômeno da concordância em tempo aparente. A amostra de Moçambique, por outro lado, até o desenvolvimento da pesquisa de Pissurno (2017), dispunha somente dos informantes constituintes da amostra básica para a investigação.

<sup>2</sup> Para maiores informações e detalhamento da distribuição das línguas autóctones em Maputo, cf. capítulo 3 do presente volume.

base o depoimento dos próprios informantes que se declaram (i) no caso de São Tomé, como falantes apenas de Português; de Português e, eventualmente, de um crioulo; ou, ainda, de um crioulo em casa e o Português nas situações fora de casa; (ii) no caso de Moçambique, como falantes de Português, que apenas compreendem pouco das línguas locais; falantes fluentes de Português e de línguas locais; ou falantes que usam mais línguas locais do que o Português. Ainda na amostra moçambicana, controlou-se, também de acordo com os depoimentos dos indivíduos, o fato de considerarem-se falantes de Português como primeira língua (L1) ou segunda (L2).

Em relação aos fatores de natureza linguística, observou-se, na análise de ambas as variedades, a possível influência dos seguintes grupos: a posição do sujeito em relação ao verbo, a distância entre o núcleo do SN e o verbo, a presença de elementos intervenientes, a configuração morfossintática do sujeito, o número de constituintes do sujeito, o paralelismo no nível oracional (do sujeito para o verbo), a animacidade do referente sujeito, a saliência fônica, o tempo/modo verbal e a transitividade verbal.

A análise desses contextos de natureza linguística e extralinguística nas duas amostras serão apresentados a seguir. Na Seção 4, relatam-se os resultados da variedade moçambicana, enquanto a Seção 5 resume os índices já observados na variedade são-tomense. A Seção 6 trará uma breve comparação entre os resultados verificados nas duas amostras.

#### 4. A CONCORDÂNCIA VERBAL NO PM

Pissurno (2017) atesta, como resultado geral, forte preferência pela concordância padrão na variedade moçambicana, revelando, consoante Labov (2003), o comportamento de uma regra semicategórica (96,8%), na qual dos 2353 dados obtidos, 2278 deles são de concordância expressa, como se vê na tabela a seguir:

**Tabela 1** Distribuição dos dados com e sem marca de P6 no PM

| CONCORDÂNCIA VERBAL DE P6 | OCORRÊNCIAS | PERCENTUAL |
|---------------------------|-------------|------------|
| Marcação de plural        | 2278/2353   | 96,8%      |
| Não marcação de plural    | 75/2353     | 3,2%       |

Fonte: Adaptado de Pissurno (2017: 155)

A análise estatística das ocorrências revelou a relevância efetiva de seis das variáveis decodificadas para os dados sem marcação plural; são elas: *língua(s) dominada(s) pelo informante, posição do sujeito, saliência fônica, escolaridade, paralelismo clausal e tipo de verbo*. Elas foram selecionadas exatamente nessa

ordem; porém, para fins de organização do texto, serão apresentadas primeiramente as variáveis sociais e, em seguida, as linguísticas.

#### 4.1 Variáveis sociais

##### a) *Língua(s) dominada(s) pelo informante*

A primeira variável, selecionada pelo programa estatístico, demonstra-se um condicionamento extremamente relevante na variedade moçambicana, não somente por sua atuação em particular, mas também em relação a outras variáveis também investigadas (como a escolaridade e o uso de Português como L1 ou L2).

Como a Tabela 2 demonstra, o maior ou menor uso das línguas autóctones – ao menos como declarado pelos próprios informantes – realmente afeta os índices de marcação de plural, como a hipótese previa, uma vez que a comparação entre os informantes que “falam mais línguas locais do que Português” e os demais confirma essa tendência. Desse modo, a baixa produtividade no uso da Língua Portuguesa nas interações sociais constitui um fator efetivamente desfavorecedor da concordância (apenas .11 para a concordância padrão).

**Tabela 2** Aplicação da marca de P6 segundo a *língua(s) dominada(s) pelo informante* no PM

| USO DAS LÍNGUAS  | OCORRÊNCIAS | PERCENTUAL | PESO RELATIVO |
|--|-------------|------------|---------------|
| Só fala Português ou apenas compreende línguas locais                      | 796/820     | 97,1%      | .49           |
| Fala, fluentemente, Português e línguas locais (em determinados contextos) | 1448/1487   | 97,4%      | .53           |
| Fala mais línguas locais do que Português                                  | 32/44       | 72,7%      | .11           |

Fonte: Adaptado de Pissurno (2017: 156)

Por outro lado, observe-se que os fatores “só fala Português e/ou apenas compreende (um pouco) as línguas locais” e “fala fluentemente línguas locais e Português” indicam comportamento bastante semelhante (apenas 4 pontos relativos de diferença), revelando uma polarização dos resultados em relação à concordância, como explica Pissurno (2017):

De um lado, observam-se os informantes que falam apenas Português (ou que podem ser considerados bilíngues passivos) juntamente com os falantes bilíngues, mostrando tendências de favorecimento à concordância verbal, ainda que seja possível notar oscilação no uso das variantes das duas categorias de falantes. De outro lado, verifica-se o favorecimento à não concor-

dância por parte apenas daquele que tem maior contato com as línguas locais do que com o Português. (PISSURNO, 2017, p. 157)

Em poucas palavras, pode-se dizer que o maior contato somente com o Português não parece ser uma condição para altos índices de marcação (.49 e .53, respectivamente), mas é, sem dúvida, um fator que distancia, em cerca de 40 pontos, esses falantes daqueles que têm maior contato com as línguas autóctones.

É válido ressaltar que, a princípio, a hipótese previa um comportamento mais escalar, demonstrando que quanto maior fosse o uso de Português, maiores seriam os índices de marcação de número. No entanto, constatou-se que o cenário de multilinguismo propicia um quadro mais complexo e um pouco diferente daquele que era esperado, especialmente porque, ao que tudo indica, ter o Português como L1 e alta escolaridade não significa, necessariamente, um domínio pleno do padrão esperado – o PE. Pissurno (2017) resume a situação:

Considera-se, nesse caso, especialmente o primeiro grupo de falantes, que contém indivíduos que afirmam “compreender línguas locais”, fato que, de acordo com Romaine (1995), já os classifica como bilíngues passivos, suscetíveis aos mesmos fenômenos variáveis de qualquer outro falante. [...] Da mesma forma, no que se refere ao fato de os indivíduos do segundo grupo falarem fluentemente línguas locais e Português, é possível que haja a confluência de muitos falantes bilíngues, aqueles que adquiriram as línguas ao mesmo tempo, desde crianças, independentemente de seu grau de escolaridade. Nesse caso, a aquisição simultânea possibilitaria um nível de proficiência muito similar nessas línguas, o que poderia causar oscilação em sua fala e, ao mesmo tempo, permitir que a aquisição das regras da Língua Portuguesa seja mais natural, resultando no leve favorecimento de uso das regras padrão. (PISSURNO, 2017, p. 157-158)

Considerando, portanto, que, em sociedades multilíngues como a moçambicana, o bilinguismo é praticamente uma regra, independentemente do nível de proficiência que o sujeito tenha nesta ou naquela língua, entende-se que monitorar a relação destes falantes com o seu conhecimento de Português como língua materna (L1) ou como segunda língua (L2) é extremamente relevante.

Contudo, vale ressaltar que até mesmo definir o estatuto da língua que usam – se L1 ou L2 – constitui uma tarefa árdua para esses indivíduos. Assim, de acordo com a proposta de Romaine (1995), acredita-se que a língua que o indivíduo indica como sendo a sua materna é aquela com a qual ele mais se identifica, em termos culturais, e não aquela que ele aprendeu primeiro. No caso de Moçambique, portanto, são identificados dois perfis distintos: de um lado, o grupo de informantes que defendem o uso exclusivo do Português, pois desejam ascensão social e melhores condições de vida e supõem que usar línguas autóctones pode atrasar seu desenvolvimento social; de outro, aqueles que acreditam no valor cultural que os idiomas nacionais

carregam e que incentivam o uso desses nos ambientes de maior representatividade (igreja, em casa, com os amigos), creditando maior importância às línguas locais do que ao Português, que seria apenas mero instrumento de comunicação geral.

Em outras palavras, a relação do indivíduo com o Português como L1 ou L2 mostra-se um grupo de fatores relevante ao aprofundar os conhecimentos sobre a variedade moçambicana. Entretanto, sua correlação com o domínio de outras línguas, como dito no início desta seção, representa um caráter bastante particular da variedade em questão, que precisa ser investigado. Em linhas gerais, Pissurno (2017) traça os resultados obtidos com a análise referente ao suposto uso que os entrevistados fazem da Língua Portuguesa:

Dentre os 18 informantes da amostra, os 11 que se declaram falantes de Português como língua materna são, especialmente, nascidos e criados em Maputo, onde a Língua Portuguesa já é majoritária e o incentivo a seu uso é extremamente valorizado. Além disso, estes também são representados por aqueles moçambicanos que tiveram maior convivência com portugueses, seja na família ou por ter passado algum tempo em Portugal e, com toda certeza, aqueles com níveis mais altos de escolaridade. [...] Considerando, de outro lado, os sete informantes da amostra que se declaram falantes de Português L2, vemos que, para muitos deles, aprender a Língua Portuguesa foi quase uma “obrigação”, algo que lhes permitiria frequentar a escola. (PISSURNO, 2017, p. 160)

A partir das considerações feitas pela autora, reflete-se sobre a atuação de uma segunda variável que envolve a questão do domínio das línguas, o grau de instrução do indivíduo. Ao que parece, ter Português como língua materna é condição para promover acesso ao conhecimento e a níveis mais elevados de escolaridade, mesmo que o domínio simultâneo de línguas autóctones influencie, em alguma medida, a variedade de Português que o indivíduo usa, porque, nesse contexto de bilinguismo, a variação em relação ao padrão tende a ser inevitável.

#### *b) Escolaridade*

Como é possível verificar na Tabela 3, o nível de escolaridade mostra-se, igualmente, um condicionamento extralinguístico bastante relevante para a marcação de número no PM. Como a hipótese indicava, confirmou-se uma escalari-idade nos índices de aplicação da regra, proporcional ao aumento dos anos em que o indivíduo esteve em contato com a educação formal.

Os informantes de ensino fundamental destoam dos demais no que se refere ao expressivo desfavorecimento da concordância verbal (.29), enquanto aqueles que possuem média ou alta escolaridade apresentam índices (.58 e .65, respectivamente) em favor da marcação. Observe-se, por outro lado, que os resultados revelam somente “16 ocorrências de ausência de marca de número no ensino

médio e 13 no ensino superior, o que nos mostra que, mesmo sendo ínfima a quantidade de “desvios”, é válido reparar a diferença de apenas três dados entre esses dois níveis de instrução escolar” (PISSURNO, 2016, p. 164). De forma geral, verifica-se que ambos os níveis de escolaridade apresentam percentuais próximos aos 98% de marcação padrão, confirmando que a diferença de comportamento entre os dois é bem pequena.

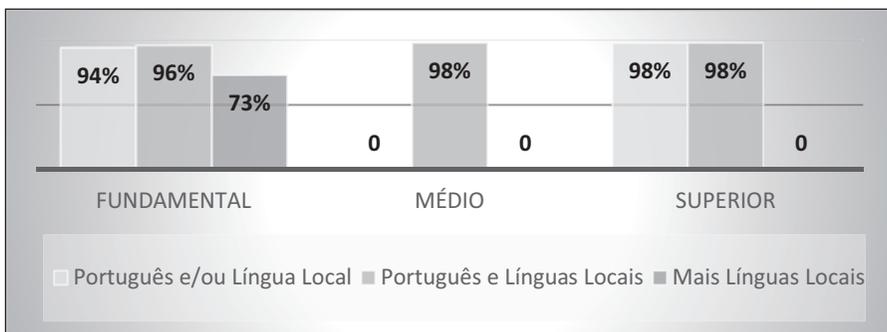
**Tabela 3** Aplicação da marca de P6 segundo a *escolaridade* no PM

| ESCOLARIDADE       | OCORRÊNCIAS | PERCENTUAL | PESO RELATIVO |
|--------------------|-------------|------------|---------------|
| Ensino Fundamental | 719/765     | 94%        | .29           |
| Ensino Médio       | 828/844     | 98,1%      | .58           |
| Ensino Superior    | 729/742     | 98,2%      | .65           |

Fonte: Adaptado de Pissurno (2017: 163)

Ao que tudo indica, os resultados observados a partir da análise da variável escolaridade parecem estar diretamente interligados ao maior ou menor uso de outras línguas, além do Português, em Moçambique. Na realidade, nessa variedade, ter mais anos de escolaridade significa maior contato com a Língua Portuguesa, já que é a língua majoritária nas escolas; assim, proporcionalmente, os informantes com menos anos de educação formal são aqueles que supostamente recebem maiores influências das outras línguas faladas na comunidade.

Com o intuito de compreender melhor essa interação entre a escolaridade e o uso de línguas, Pissurno (2017) realizou um cruzamento entre as duas variáveis, de modo a aferir como o contato entre línguas pode afetar tanto os informantes que se declaram falantes de Português como L1, quanto aqueles que se afirmam bilíngues. A Figura 1 apresenta os resultados de tal cruzamento:



**Figura 1** Cruzamento de dados com marca de P6: *língua(s) dominada(s) pelo informante x escolaridade* no PM

Fonte: Adaptado de Pissurno (2017: 164)

Como a Figura 1 revela, apenas o contexto do ensino fundamental em Moçambique apresenta resultados para os três fatores estabelecidos para o uso de línguas na região. Em outras palavras, somente esse grupo de indivíduos com menor escolaridade declara usar as línguas locais com maior ou menor proficiência e produtividade. Contudo, é importante destacar que o único índice abaixo dos 90% de concordância padrão é o encontrado para o fator “fala mais línguas locais do que Português”. Esse baixo percentual de realização da marca de plural, em apenas 73% das ocorrências, é referente à fala do único informante de toda a amostra que se declarou como tal tipo de falante. Em poucas palavras, trata-se de um indivíduo bastante diferente dos outros perfis entrevistados, por conta de sua baixa escolaridade e suas notáveis dificuldades com a Língua Portuguesa de forma geral<sup>3</sup>.

Em relação ao ensino médio, constata-se ser esse o único fator que apresenta apenas falantes bilíngues. Por fim, no ensino superior, os falantes são ora falantes apenas de Português, ora de Português e de línguas locais simultaneamente, porém ambos os fatores indicam elevada preferência pela concordância (98%).

Do exposto, reforça-se a hipótese de que ter contato com línguas autóctones influencia, de fato, a marcação de pluralidade no PM, representando as maiores tendências à ausência de marcas quanto maior for o uso desses idiomas. Além disso, como aponta Pissurno (2017):

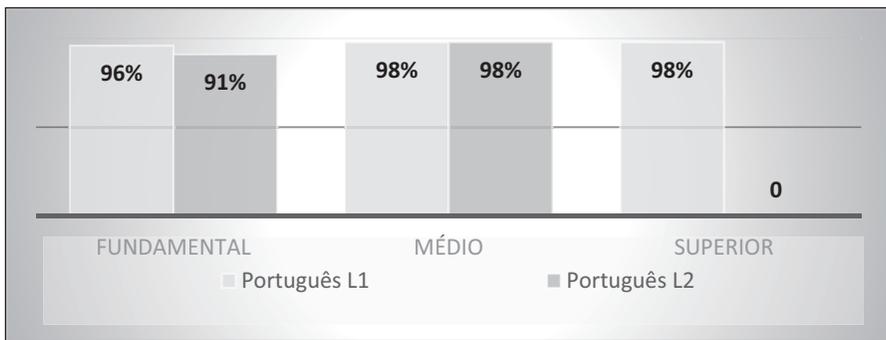
Observa-se que ser bilíngue é uma realidade muito presente nessa amostra, independentemente do nível de escolaridade. Com isso, nota-se que, para a maioria dos informantes aqui analisados, o Português foi adquirido com algum tipo de influência das línguas locais, mesmo que alguns afirmem o contrário. Às vezes, o simples fato de alguém da família ser fluente em idiomas locais já pode exercer algum tipo de influência no Português falado por eles. (PISSURNO, 2017, p. 165)

Por fim, destaca-se, ainda, que, de acordo com Pissurno (2017), sendo a escolaridade uma variável tão importante para a configuração da regra de concordância verbal no Português de Moçambique, é válido considerar que os professores de Língua Portuguesa passam pelas mesmas interferências linguísticas que

---

<sup>3</sup> Como esclarece Pissurno (2017, p. 158): “Esse informante (PMOC1H) possui apenas a 4ª classe do ensino fundamental, ou seja, seu contato com a Língua Portuguesa foi, de fato, muito menor do que o de alguns outros informantes pertencentes aos outros grupos. Seus resultados referentes à marcação de concordância chegam aos 73% (limiar referente à variação) e ele apresenta dificuldades em outras áreas relacionadas ao processamento do mecanismo de concordância em geral, como a de gênero e a nominal, por exemplo, que ele utiliza pouquíssimas vezes de acordo com as regras do Português padrão ao longo de sua entrevista.”

seus alunos e, conseqüentemente, transmitem regras de um Português (seja como L2 ou como uma L1 com influência, inerente, das línguas autóctones) que podem não refletir totalmente aquelas que são consideradas padrão no Português Europeu, dando margem a contextos variáveis. Assim, reforça-se a existência de uma conexão muito forte entre a escolaridade e o uso de Português nessa sociedade. Como a Figura 2 revela, “não existe qualquer informante de ensino superior que seja falante de Português como L2” e, ao mesmo tempo, o menor índice de marcação de plural (91%) “é observado exatamente nos indivíduos de ensino fundamental que não têm o Português como sua língua materna” (PISSURNO, 2017, p. 166.)



**Figura 2** Cruzamento de dados com marca P6: *língua materna x escolaridade no PM*

Fonte: Adaptado de Pissurno (2017: 166)

Em suma, conclui-se que a escolaridade e as línguas dominadas pelos indivíduos dessa amostra são, sem dúvida, características extremamente relevantes para o entendimento da variedade moçambicana de Português. No entanto, elas devem ser analisadas em relação a outras variáveis extralinguísticas que, apesar de não serem estaticamente relevantes, demonstram a atuação de vários fatores particulares à realidade social da variedade em questão. Devido à delimitação do presente capítulo, optou-se por apresentar apenas a influência do Português como L1 ou L2 nos dados<sup>4</sup> obtidos.

## 4.2 Variáveis linguísticas

### a) *Posição do sujeito*

Conforme a Tabela 4, pode-se notar o efeito inegável da posição do sujeito no condicionamento das marcas de número no PM, já que nos contextos de sujei-

<sup>4</sup> Pissurno (2017) ainda levanta questionamentos relacionados a outras variáveis como a faixa etária e as características individuais dos informantes.

to posposto a tendência é a de desfavorecimento da concordância padrão (.17), enquanto o sujeito anteposto favorece o uso das marcas (.58).

**Tabela 4** Aplicação da marca de P6 segundo a *posição do sujeito* no PM

| POSIÇÃO DO SUJEITO | OCORRÊNCIAS | PERCENTUAL | PESO RELATIVO |
|--------------------|-------------|------------|---------------|
| Sujeito anteposto  | 1196/1228   | 97,4%      | .58           |
| Sujeito posposto   | 202/223     | 90,6%      | .17           |

Fonte: Adaptado de Pissurno (2017: 175)

Seguem abaixo alguns exemplos encontrados na amostra analisada, que confirmam a tendência de não haver marcas quando o sujeito aparece depois do verbo (exemplos 1 e 2), diferentemente do que ocorre quando o sujeito se encontra antes da forma verbal (exemplos 3 e 4):

- (1) tinha velhos colono ainda... saiu os colono consolidar a vida (PMOC1H)
- (2) às vezes vinha dois três... conforme eles iam crescendo não é... (PMOC3M)
- (3) as cidades começam a ficar um pouco mais aceleradas e começam a surgir de certa maneira neste momento de certa maneira focos né criminalidade (PMOB3H)
- (4) eles fazem essas coisas todas na praia (PMOA2H)

Ainda sobre a influência do sujeito sobre os resultados encontrados, é importante destacar a atuação de uma configuração morfossintática específica desse constituinte, quando este conta com a presença de pronome relativo. Dos 32 dados sem marca de número em contexto de sujeito anteposto encontrados, 10 eram de sujeito representado pelo relativo *que*. Esse comportamento sugere que a presença do pronome relativo exerce influência sobre a marcação de pluralidade verbal, comportando-se, assim, como contexto desfavorecedor da concordância, conforme se pode atestar nos exemplos a seguir:

- (5) para nossas criança elas que vai fazer furo / futuro melhora... (PMOC1H)
- (6) ah: é normal colegas meus que tá no outro no outro semestre da faculdade escreverem muito mal português... (PMOA3H)
- (7) conflitos armados e: : guerra que assola lá na zona centro... (PMOB1H)
- (8) casamentos daqui não dura até pior que os que casa oficialmente é normal uma pessoa casar este ano e próximo ano já divorciou (PMOB1M)

A respeito dos exemplos destacados, Pissurno (2017) faz duas observações relevantes sobre esse condicionamento:

- (i) apenas dois desses dados (5 e 6) não são construídos com verbos representativos do grau 1 de saliência, ou seja, a maioria dos casos de ausência de marcas de plural em sujeitos representados pelo pronome *que* são, coincidentemente, verbos de baixa saliência fônica, que por si só, já são favorecedores da não marcação.
- (ii) apenas um foi produzido por um informante que não fosse do ensino fundamental [...] mesmo que seja apenas um exemplo do ensino superior, pode-se ver que o pronome relativo pode atuar, ainda que muito discretamente, no conjunto de fatores, para a ausência de concordância. No caso desse informante, essa foi a única ocorrência de ausência de marca de número no seu discurso (6). (PISSURNO, 2017, p. 178)

### b) Saliência fônica

O comportamento da variável saliência fônica (Tabela 5) confirma as hipóteses previstas de que os graus (1 e 2) de menor distanciação fônica entre as formas de singular e plural desfavoreceriam a concordância (.38 e .37, respectivamente), enquanto os graus (3, 4 e 5) de maior saliência favoreceriam a presença das marcas.

**Tabela 5** Distribuição dos dados com marca de P6 segundo *saliência fônica* no PM

| SALIÊNCIA FÔNICA           | OCORRÊNCIAS | PERCENTUAL | PESO RELATIVO |
|----------------------------|-------------|------------|---------------|
| Grau 1 – come(m) / fala(m) | 1268/1319   | 96,1%      | .38           |
| Grau 2 – faz/fazem         | 196/201     | 97,5%      | .37           |
| Grau 3 – dá/dão            | 331/339     | 97,6%      | .54           |
| Grau 4 – comeu/comeram     | 145/237     | 96,2%      | .56           |
| Grau 5 – é/são             | 255/257     | 99,2%      | .94           |

Fonte: Adaptado de Pissurno (2017: 180)

Observe-se, ainda, que, enquanto os graus 3 e 4 demonstram comportamento similar quanto ao favorecimento da concordância padrão (.54 e .56, respectivamente), no grau 5 a probabilidade de favorecer a marcação é altíssima (.94), o que corresponde a um subconjunto de dados, em que a ausência de marcas de número ocorreu em apenas dois dados. A seguir, apresenta-se um exemplo de verbo para cada um dos graus de saliência fônica estabelecido:

#### Grau 1

- (9) que eles *fica* devendo em dinheiro quando às vezes nem tem dinheiro... (PMOB2H)

**Grau 2**

- (10) os mais antigos faz muita confusão na profissão polícia como é possível pagar e eles pronto não pegam a multa deixa seguir e vai andando (PMOC1M)

**Grau 3**

- (11) porque já tem encontrado alguns professores que dão teste então dão teste em troca de favores (PMOA1M)

**Grau 4**

- (12) algumas pessoas saíram ali e disseram que... nós não fomos atacados pelos homens da RENAMO (PMOA2H)

**Grau 5**

- (13) mas algumas crianças são boas... tem um bom aproveitamento... comportam-se bem... (PMOC2M)

Cabe ressaltar que o comportamento do PM em relação à saliência fônica mostra o quanto essa variável é importante para a caracterização das variedades do Português, como já destacaram Vieira; Brandão (2014) ao afirmarem que PB e PE constituem sistemas opostos de concordância, porque a ausência da mesma é condicionada por fatores diferentes. Dentre esses fatores, as autoras destacam a saliência, variável relevante em amostras de variedades brasileiras, mas não na variedade europeia. Scherre; Naro (2006) corroboram esse argumento, quando apontam que a saliência não é significativa quando produzida por falantes altamente escolarizados. Dessa forma, Pissurno declara:

a preferência exclusiva pela variante padrão no PE acaba por apagar a atuação desse grupo de fatores, diferentemente do que ocorre no PB, sobretudo nas variedades populares, e no PM, que aparentemente têm, em termos qualitativos (embora com padrões quantitativos bem diferentes), comportamento similar no fenômeno da concordância. (PISSURNO, 2017, p. 183)

Em poucas palavras, entende-se que os condicionamentos que afetam a concordância no PM, principalmente a escolaridade e a saliência fônica, indicam – não obstante os diferentes índices gerais de não concordância – semelhanças relativas aos contextos de ausência de marca que ocorrem nas variedades brasileiras, distanciando, de certa forma, a variedade moçambicana, em termos qualitativos, de seu modelo europeu de aprendizagem.

c) *Paralelismo oracional*

A partir do princípio geral do paralelismo, de que marcas levam a marcas e zeros levam a zeros (SCHERRE; NARO, 1993), constata-se, nos dados da variedade em questão, a confirmação dessa hipótese.

Como se vê na Tabela 6, dados em que o verbo é antecedido de um sintagma com marcas favorecem a concordância (.52), ao passo que os sintagmas sem marca acarretam a ausência da mesma nos verbos (.5 para a concordância padrão).

**Tabela 6** Aplicação da marca de P6 segundo *paralelismo clausal* no PM

| PARALELISMO ORACIONAL   | OCORRÊNCIAS | PERCENTUAL | PESO RELATIVO |
|-------------------------|-------------|------------|---------------|
| SN com marcas de plural | 1168/1194   | 97,8%      | .52           |
| SN sem marcas de plural | 4/8         | 50%        | .5            |
| SPrep marcado ou não    | 27/29       | 93,1%      | .23           |
| SN's coordenados        | 9/11        | 81,8%      | .12           |

Fonte: Adaptado de Pissurno (2017: 184)

Os exemplos a seguir confirmam os índices obtidos: enquanto no (14) a presença de marcas explícitas de plural no sintagma “as meninas” acarreta a marcação padrão do verbo “vivem”, no (15) a ausência delas em “as mulher” gera o apagamento da marca também no verbo “deixa”:

(14) as meninas *vivem* ali... vêm dos distritos... são meninas de famílias... muito carentes... (PMOC3M)

(15) porque as mulher daqui não *deixa* a possibilidade (PMOA1H)

Sobre a atuação dos outros contextos em relação ao desfavorecimento da concordância no PM, Pissurno (2017) esclarece:

o cancelamento das marcas de pluralidade, de fato, diminui à medida que os SN's sujeitos apresentam quaisquer marcas de número. Sendo assim, a tendência de não marcação nos contextos analisados obedeceria a seguinte ordem: SN sem marcas > SN's coordenados > SPrep marcado ou não > SN com marcas. (PISSURNO, 2017, p. 185)

Assim, pode-se concluir que os resultados encontrados confirmam as tendências gerais sobre o paralelismo: a ausência de marcas formais no sujeito enunciado favorece a ausência de concordância padrão nos verbos que o seguem.

#### d) *Tipo de verbo*

Por fim, a última variável selecionada como relevante para o PM foi o tipo de verbo. Observem-se os resultados obtidos na Tabela 7, a seguir:

**Tabela 7** Aplicação de marca de P6 segundo *tipo de verbo* no PM

| TIPO DE VERBO              | OCORRÊNCIAS | PERCENTUAL | PESO RELATIVO |
|----------------------------|-------------|------------|---------------|
| Inergativos e Inacusativos | 295/319     | 92,5%      | .37           |
| Transitivos                | 1549/1589   | 97,5%      | .56           |
| Copulativos                | 434/445     | 97,5%      | .40           |

Fonte: Adaptado de Pissurno (2017: 186)

De acordo com esses resultados, verifica-se que os verbos transitivos são os favorecedores da marcação de plural (.56) na presente amostra. Por outro lado, os contextos desfavorecedores são os verbos inergativos/inacusativos (.37) – como no exemplo (16) – e os copulativos (.40), como no exemplo (17):

(16) então ainda *existe* muitos lobolos (PMOA1M)

(17) e os donos da casa também que já *era* velhinhos (PMOB2M)

Ao considerar a variável *tipo de verbo*, não se pode deixar de lembrar que existe uma correlação muito forte desse grupo de fatores com a posição do sujeito em questão, como apontam Scherre; Naro; Cardoso (2007). No referido trabalho sobre a amostra moçambicana, tal correlação foi confirmada<sup>5</sup>, tendo tido atuação mais forte a posição do sujeito em relação ao verbo. De acordo com os referidos autores, “independentemente do tipo de verbo, qualquer argumento ou sintagma à direita do verbo tende, relativamente, a diminuir as marcas de concordância explícita” (SCHERRE; NARO; CARDOSO, 2007, p. 312).

A posição do sujeito constitui, portanto, um condicionamento que se pode considerar “geral” da Língua Portuguesa, pois ocorre tanto nas variedades africanas quanto na brasileira e até mesmo na europeia, como demonstram Vieira; Bazenga (2013). As autoras ainda afirmam que esse grupo de fatores “altera padrões de concordância em línguas de perfis e genealogias muito diferentes”, o que “faz supor que não se trate de efeito efetivamente variável nesse caso, mas de comportamento estrutural gramaticalmente diferenciado” (VIEIRA; BAZENGA, 2013, p. 55).

Em síntese, as tendências encontradas demonstram que favorecem a concordância no Português de Moçambique: (i) das variáveis extralinguísticas: níveis

<sup>5</sup> No desenvolvimento das variáveis linguísticas na dissertação, Pissurno (2017) apresenta os cruzamentos realizados entre as variáveis para a confirmação de tais hipóteses. Além disso, a autora apresenta outros testes executados sobre as variáveis em debate, que não cabiam no escopo do presente capítulo.

médio e superior de escolaridade e falantes com mais contato com Língua Portuguesa e/ou bilíngues; (ii) das variáveis linguísticas em relação ao sintagma nominal: os sujeitos em posição anterior ao verbo e com marcas explícitas de plural; (iii) e das linguísticas em relação ao sintagma verbal, os verbos transitivos e com maior grau de saliência fônica (graus 3, 4 e 5). Por outro lado, baixa escolaridade, maior contato exclusivo com línguas autóctones, sujeitos pospostos, sintagmas sem marcas de plural, verbos inergativos/inacusativos e copulativos, assim como aqueles com baixa saliência fônica entre as formas singular e plural são os contextos que desfavorecem as marcas de concordância na amostra em estudo.

Em poucas palavras, resume-se a situação de Moçambique nas considerações de Pissurno (2017):

[...] pode-se concluir que os 75 dados de não concordância coletados são bastante variados, não sendo de uma mesma natureza. Nesse sentido, fica viável uma abordagem contrastiva, não só dos índices quantitativos, mas também da natureza dos dados da concordância na variedade moçambicana do Português, em termos qualitativos, com os padrões descritos para o PB e o PE. Embora quantitativamente os dados do PM estejam mais próximos dos registrados para o PE (como uma regra semicategórica) do que para o PB (como uma regra variável), a natureza estrutural dessas ocorrências e o comportamento dos indivíduos, caracterizados em termos sociolinguísticos, ao que parece, aproximariam, qualitativamente, o PM do PB e não do PE. (PISSURNO, 2017, p. 195-196)

Resta saber se tal comportamento também ocorre em outras variedades africanas que sofrem influências semelhantes em relação ao multilinguismo. Na próxima seção, portanto, serão relatados os resultados da variedade de São Tomé.

## 5. A CONCORDÂNCIA NO PST

No âmbito do projeto *Estudo comparado dos padrões de concordância em variedades africanas, brasileiras e europeias do Português*, o Português de São Tomé e Príncipe vêm sendo examinado com o objetivo de verificar os traços específicos da variedade que permitam, não só sua caracterização, como também sua comparação com outras variedades de Português, especialmente em relação ao fenômeno da concordância nominal e verbal.

A partir de tais objetivos, Brandão; Vieira (2012), em uma investigação que considerava 22 informantes da amostra são-tomense (17 informantes da amostra básica e 5 informantes da amostra complementar de ensino médio e superior), já atestavam a forte preferência pela concordância verbal (93,1% de marcação) nessa variedade. Por sua vez, Vieira; Bazenga (2013, 2015), em uma análise contras-

tiva das variedades do Português, observaram o comportamento da amostra básica (17 informantes), relatando, igualmente, altos índices de favorecimento das marcas de número (92,1% de concordância padrão). Assim, pode-se dizer que, de maneira geral, atua sobre essa variedade uma regra variável de concordância verbal.

Com base nesses trabalhos, Vieira; Silva (2017) examinaram, além das entrevistas da amostra básica, uma amostra complementar de 10 informantes jovens (10 a 19 anos), e encontraram 100 ocorrências com ausência de marcas de número dentre os 1053 dados obtidos. Dessa forma, obteve-se 90,5% de preferência pela concordância verbal padrão, conferindo, novamente, um estatuto de regra variável para o fenômeno na variedade em questão, semelhante aos resultados encontrados nos trabalhos supracitados.

**Tabela 8** Distribuição dos dados com e sem marca de P6 no PST

| CONCORDÂNCIA VERBAL DE P6 | PERCENTUAL | OCORRÊNCIAS |
|---------------------------|------------|-------------|
| Marcação de plural        | 90,5%      | 953/1053    |
| Não marcação de plural    | 9,5%       | 100/1053    |

Fonte: Vieira; Silva (2017: 22)

As variáveis com comportamento relevante para o tratamento estatístico da amostra de dados são-tomenses foram<sup>6</sup>, nesta ordem, escolaridade, animacidade do referente sujeito, posição do sujeito, sexo e paralelismo clausal. Assim como se fez para a descrição dos resultados do PM, as variáveis sociais serão apresentadas antes das linguísticas.

## 5.1 Variáveis sociais

### a) Escolaridade

Semelhantemente ao que ocorre em Moçambique, em São Tomé os informantes de ensino fundamental mostram maior tendência ao desfavorecimento da

<sup>6</sup> Destaca-se que, em Brandão; Vieira (2012), além das variáveis destacadas como relevantes por Vieira; Silva (2017), as autoras também obtiveram seleção estatística das variáveis *faixa etária* e *saliência fônica* na amostra que analisaram. Por sua vez, Vieira; Bazenga (2015) relacionam como variáveis significativas para a amostra o *paralelismo*, a *posição do sujeito*, a *escolaridade*, o *sexo* e o *traço semântico do sujeito*. Portanto, observa-se que, apesar de apresentarem algumas diferenças, nas três investigações há semelhanças quanto ao comportamento das variáveis *escolaridade*, *sexo* e *posição do sujeito*, como condicionadores do fenômeno.

concordância verbal (.25), enquanto os que possuem ensino superior apresentam índice elevado de preferência pelas marcas de número (.82). Os informantes de ensino médio apresentam um comportamento intermediário em relação à marcação de pluralidade (.48).

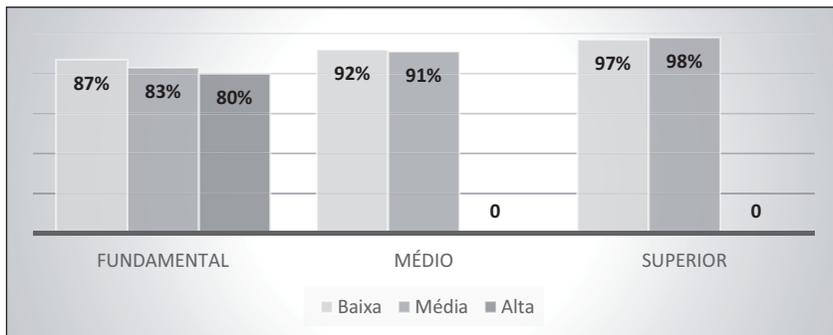
**Tabela 9** Aplicação da marca de P6 segundo a *escolaridade* no PST

| ESCOLARIDADE | OCORRÊNCIAS | PERCENTUAL | PESO RELATIVO |
|--------------|-------------|------------|---------------|
| Fundamental  | 290/348     | 83%        | .25           |
| Médio        | 371/406     | 91%        | .48           |
| Superior     | 292/299     | 98%        | .82           |

Fonte: Vieira; Silva (2017: 22)

Em termos percentuais, observa-se a mesma escolaridade de marcação que as variedades moçambicana e brasileira, revelando que os índices mais altos de marcação de número se realizam na fala dos informantes que possuem mais anos de educação formal. Por outro lado, menos anos de escolaridade registram menores índices de marcação de plural, o que indica não só menor domínio das regras do Português “padrão”, como também maior contato com as línguas crioulas faladas no país.

A partir de tal constatação, Vieira; Silva (2017) propõem um cruzamento entre as duas variáveis, resumido na Figura 3, que segue:



**Figura 3** Cruzamento de dados com marca P6: *língua(s) de intercomunicação* x *escolaridade* no PST

Fonte: Adaptado de Vieira; Silva (2017: 23)

De acordo com as autoras,

Valendo-se da proposta de Labov (2003) – que propõe a categorização de regras em categóricas (100% de uma das formas), semicategóricas (95 a 99% de uma das formas) e variáveis (5-95% de uma das formas) –, observa-se que os universitários apresentam cerca de 98% de concordância, enquadrando-se na classificação de uma regra semicategórica. Os falantes de ensino médio e fundamental apresentam variação, com evidente diferença entre eles. É importante destacar que o cruzamento com a variável “língua de comunicação” se mostra importante para a compreensão desse quadro: os falantes de ensino fundamental que usam o crioulo em casa são os que mais apresentam ausência de concordância, exatamente ao usarem o Português, fora de casa. (VIEIRA; SILVA, 2017, p. 10)

Em poucas palavras, as evidências encontradas no cruzamento das duas variáveis confirmam que, na sociedade são-tomense, como sinalizado nos trabalhos anteriores, mais anos de escolaridade pressupõem maior contato com a Língua Portuguesa, acarretando índices mais elevados de marcação de número, assim como afirmam Brandão; Vieira (2012):

De fato, o prestígio que recebe o Português em terras são-tomenses e o maior ou menor propósito de dominar essa língua, o que pode se tornar possível graças ao acesso à escolaridade, trazem por consequência o maior ou menor domínio das regras canônicas de concordância. (BRANDÃO; VIEIRA, 2012, p. 1055)

#### b) *Sexo*

Em relação à variável sexo, verifica-se que as mulheres são, de fato, as maiores favorecedoras das marcas (.64), o que se correlaciona com o pressuposto laboviano de que elas tenderiam a ser as propulsoras das variantes de maior prestígio. Os homens, por sua vez, revelam um comportamento de maior uso da variante sem marcas de número (.40 para a concordância).

**Tabela 10** Aplicação da marca de P6 segundo o *sexo* no PST

| SEXO DO INFORMANTE | OCORRÊNCIAS | PERCENTUAL | PESO RELATIVO |
|--------------------|-------------|------------|---------------|
| Homem              | 508/579     | 88%        | .40           |
| Mulher             | 445/474     | 94%        | .64           |

Fonte: Vieira; Silva (2017: 23)

Como se vê, esse resultado reforça os relatos discutidos por Vieira; Bazenga (2015), confirmando as tendências ora encontradas em relação ao sexo do informante:

[...] quanto maior a inserção da mulher na vida social, econômica e profissional (como parece ser o caso de Copacabana e São Tomé), maior a tendência a se efetivar o uso das marcas de plural, forte sinalizador de detenção da cultura letrada. (VIEIRA; BAZENGA, 2015, p. 52)

## 5.2 Variáveis linguísticas

### a) *Animacidade*

A variável animacidade confirma a hipótese de que sujeitos com traço [-animado] tendem a influenciar a ausência de uso das marcas. Como a Tabela 11 aponta, sujeitos inanimados desfavorecem a marcação (.22 para concordância), ao passo que os animados a favorecem (.57).

**Tabela 11** Aplicação da marca de P6 segundo a *animacidade* no PST

| ANIMACIDADE DO SUJEITO | OCORRÊNCIAS | PERCENTUAL | PESO RELATIVO |
|------------------------|-------------|------------|---------------|
| [+ animado]            | 816/879     | 93%        | .57           |
| [- animado]            | 137/174     | 79%        | .22           |

Fonte: Vieira; Silva (2017: 24)

Os exemplos a seguir, retirados das entrevistas da amostra PST, confirmam essas tendências:

(18) as pessoas me *olhavam* com bons olhos (PSTA3H)

(19) e depois eu com essas dificuldades todas, mesmo assim, as dificuldades não *acaba*". (PSTA2H)

### b) *Posição do sujeito*

Quanto à posição do sujeito, confirma-se a tendência, já referida, de que sujeitos em posição anterior ao verbo favorecem a presença de marcas (.54), enquanto em posição posposta ao verbo tendem a induzir a ausência de marcação (apenas .14 para a concordância).

**Tabela 12** Aplicação da marca de P6 segundo a *posição do sujeito* no PST

| POSIÇÃO DO SUJEITO | OCORRÊNCIAS | PERCENTUAL | PESO RELATIVO |
|--------------------|-------------|------------|---------------|
| Sujeito Anteposto  | 633/688     | 92%        | .54           |
| Sujeito Posposto   | 38/53       | 72%        | .14           |

Fonte: Vieira; Silva (2017: 25)

Os exemplos a seguir demonstram tais tendências: no exemplo (20), o sujeito “as coisas”, em posição anteposta ao verbo “corram”, propicia a marcação de número padrão, enquanto em (21) o sujeito “aquelas lula grande”, em posição posposta ao verbo, provoca a ausência de marca de número no verbo “morre”:

(20) espero que as coisas corram bem (PSTB2M)

(21) não é lulinha, morre aquelas lula grande mesmo”. (PSTA1H)

c) *Paralelismo oracional*

De acordo com a hipótese inicial para a variável *paralelismo oracional*, sintagmas nominais sujeitos com marcas de número favoreceriam a presença de marcas no verbo, ao passo que a ausência delas no SN levaria à não marcação do verbo. Tal hipótese também foi confirmada na amostra em questão, como os dados revelam. No exemplo (22), o sintagma não marcado “financeira” desfavorece a presença da marca de número no verbo “permitia”, enquanto em (23), todas as marcas presentes no SN “os meus filhos” favorecem a marcação plural no verbo “passavam”:

(22) as condições financeira não *permitia* prosseguir (PSTA1H)

(23) os meus filhos *passavam* no jardins (PSTA1M)

Logo, a presença de marcas no SN favorece a marcação no verbo (.53), enquanto o SN sem marcas acarreta a ausência de concordância do verbo (.16), conforme apresentado na Tabela 13.

**Tabela 13** Aplicação da marca de P6 segundo o *paralelismo oracional* no PST

| PARALELISMO ORACIONAL   | OCORRÊNCIAS | PERCENTUAL | PESO RELATIVO |
|-------------------------|-------------|------------|---------------|
| SN com marcas de plural | 625/672     | 93%        | .53           |
| SN sem marcas de plural | 17/29       | 59%        | .16           |

Fonte: Vieira; Silva (2017: 25)

Por fim, Vieira; Silva (2017) concluem:

Em outras palavras, os padrões de concordância em PST urbano verificados no presente estudo não correspondem exatamente aos verificados para o PE (com dados semicategóricos em geral) – apesar de este ser o modelo que os são-tomenses supostamente tendem a seguir, por ser a prestigiosa variedade do colonizador –, mas também não chegam a se igualar aos das variedades brasileiras, sobretudo as populares, com índices mais baixos de realização da marca de pluralidade. (VIEIRA; SILVA, 2017, p. 17).

Resumidamente, as tendências encontradas na amostra analisada demonstram que favorecem a concordância no Português de São Tomé: (i) das variáveis extralinguísticas: os indivíduos com nível superior de escolaridade e as mulheres; (ii) das variáveis linguísticas: os sujeitos animados, antepostos ao verbo e com marcas explícitas de plural. Por outro lado, menos anos de escolaridade, sujeitos inanimados e pospostos, e sintagmas sujeitos sem marcas de plural são os contextos que desfavorecem as marcas de concordância na amostra do PST. Vale lembrar que, conforme mostram os resultados do cruzamento entre os grupos de fatores em questão, a variável relacionada ao uso de línguas crioulas demonstra correlação direta com a escolaridade, indicando que o maior contato com elas pressupõe indicação de menos anos de estudo formal e, conseqüentemente, de índices mais altos de não concordância.

Ressalte-se que os fatores sociais, ao que tudo indica, mostram-se altamente relevantes para a caracterização das variedades africanas do Português, especialmente as exploradas no presente capítulo (PST e PM), devido ao elevado prestígio associado ao uso de Português, como declaram Vieira; Brandão (2012) sobre o PST:

[...] ao que tudo indica, o valor social que se atribui à não concordância seria compatível com um estereótipo (nos termos de Labov (1972)) ou, na melhor das hipóteses, com um marcador. Por essa razão, o fenômeno variável funciona como nítido identificador do perfil social do usuário da língua. No PST – que configura um caso de variedade do português em formação, com normas objetivas ainda não totalmente definidas e assumidas pelos diferentes grupos de falantes –, sabe-se que dominar a concordância segundo a norma culta que lhe serve de modelo, a do Português Europeu, seria um dos índices de identificação de pleno domínio da Língua Portuguesa. (BRANDÃO; VIEIRA, 2012, p. 1061)

## 6. A CONCORDÂNCIA NO PM E NO PST: UMA BREVE COMPARAÇÃO

Os resultados dos trabalhos revisitados no presente capítulo permitem a verificação de tipos de regra diferentes em relação ao fenômeno da concordância verbal nas variedades estudadas. Enquanto a variedade são-tomense apresenta um comportamento um pouco mais parecido com o das variedades brasileiras urbanas, por atuar dentro dos índices de variação (90,5%), a variedade moçambicana sinaliza um desempenho semicategórico (96,8%) – nos termos de Labov (2003) – de marcação de número, mais próximo dos valores encontrados no Português Europeu. No entanto, as variáveis sociais, sobretudo no que se refere ao perfil do informante quanto ao emprego das línguas locais ou não, e os contextos qualitativos que atuam sobre a não concordância do PM também o aproximam, em alguma medida, das variedades brasileiras.

Dessa forma, pode-se compreender que, embora ambas as variedades tenham como modelo de aquisição o Português Europeu, elas apresentam uma realidade linguística que interfere diretamente no uso da norma considerada padrão. Sendo assim, atesta-se, de acordo com os fatores aqui analisados, o afastamento dos índices (semi)categorísticos de concordância verificados em amostras do PE, especialmente por conta da sensibilidade aos fatores extralinguísticos como a escolaridade, que não impõe qualquer tipo de restrição quanto ao uso das marcas pelos europeus, e o contato com outras línguas (antes, durante ou após o aprendizado de Língua Portuguesa), além de alguns fatores linguísticos que atuam de forma similar em variedades brasileiras. O quadro a seguir sintetiza as variáveis relevantes para as duas variedades africanas:

**Quadro 1** Variáveis selecionadas para a implementação da ausência de concordância de P6 por amostra (PM x PST)

| MOÇAMBIQUE (PM)                       | SÃO TOMÉ E PRÍNCIPE (PST) |
|---------------------------------------|---------------------------|
| Língua(s) dominada(s) pelo informante | Escolaridade              |
| Posição do sujeito                    | Animacidade do sujeito    |
| Saliência Fônica                      | Posição do sujeito        |
| Escolaridade                          | Sexo                      |
| Paralelismo oracional                 | Paralelismo oracional     |
| Tipo de verbo                         |                           |

Fonte: Pissurno (2017); Vieira; Silva (2017)

A partir do Quadro 1, verifica-se que atuam nas duas variedades africanas a escolaridade, a posição do sujeito e o paralelismo oracional, ainda que as duas primeiras variáveis apareçam em ordem diferente. A variedade são-tomense se particulariza em relação à seleção de animacidade do sujeito e sexo do informante, enquanto apenas na moçambicana se observa a atuação das línguas dominadas pelo informante, a saliência fônica e o tipo de verbo<sup>7</sup>.

Em geral, pode-se dizer que as variedades demonstram tendências semelhantes quanto ao uso de línguas além do Português, uma vez que, independentemen-

<sup>7</sup> Ressalte-se que Vieira; Silva (2017) descrevem a interferência entre as variáveis saliência fônica e tipo de verbo no PST (não selecionadas estatisticamente, porém indicando comportamento relevante para a não marcação), que não foram inseridas no presente capítulo por questões de delimitação de espaço.

te do nível de escolaridade, as duas representam uma realidade de contato entre línguas que atua sobre as regras de marcação plural. No caso da variedade são-tomense, observa-se esse comportamento, principalmente, na frequência média de uso de um crioulo, ao passo que na moçambicana, tal cenário é refletido naqueles que falam fluentemente o Português e uma (ou mais) línguas locais.

Ademais, os dados mostram que, nas duas variedades, é evidente a influência do uso, seja de uma língua local ou de um crioulo, nos valores obtidos e sua relação com o ensino de Português, já que ela é inversamente proporcional ao nível de escolaridade. Em outras palavras, quanto mais anos de escolaridade (ensino médio e superior), maior o contato com o Português considerado formal e, conseqüentemente, menor interferência das outras línguas, refletindo os índices mais altos de marcação plural nas duas amostras (todos acima dos 90%). Esse resultado confirma o caráter fundamental da instrução escolar na aprendizagem do Português nas variedades em questão e ainda permite a confirmação da hipótese de que o contato linguístico pode, de fato, afetar os índices de concordância verbal nas variedades de Língua Portuguesa em formação.

Assim, como concluem Vieira; Bazenga (2015) sobre a escolaridade e o uso de línguas:

A correlação entre as referidas variáveis, na qualidade de evidência empírica, favorece a hipótese de que o maior contato linguístico se correlacione à menor realização da marca, hipótese empregada nos estudos que defendem ser a não concordância brasileira, nos níveis e contextos existentes, oriunda da forte situação de contato linguístico entre negros, índios e brancos. (VIEIRA; BAZENGA, 2015, p. 51)

Compreende-se, portanto, que a realidade linguística das variedades africanas deve ser criteriosamente observada por conta da forte realidade multilinguística presente nessas sociedades, em especial na moçambicana, cujo bilinguismo/multilinguismo tem se mostrado cada vez mais frequente entre os falantes de todas as idades e origens. Em poucas palavras, vale considerar que, em menor ou maior grau, a depender de todos os fatores extralinguísticos relacionados ao uso das línguas em cada sociedade, é imperativo não tratar tais variedades de Português como isentas da influência do multilinguismo, como se elas fossem reproduções do Português Europeu e totalmente distanciadas de qualquer “desvio” associado ao uso de um crioulo ou de uma língua Banto.

## 7. CONSIDERAÇÕES FINAIS

A análise contrastiva entre a variedade são-tomense e a moçambicana do Português aqui realizada revelou diferenças em relação ao estatuto da regra de

concordância, ao constatar, aparentemente, uma regra variável na primeira – embora com alto índice de realização da concordância – e uma semicategórica na segunda. De toda forma, os condicionamentos que desfavorecem a concordância verbal em ambas as amostras demonstraram que se trata de graus de efetivação de uma regra que pode ser considerada variável nas duas variedades, sobretudo se considerada a complexidade da composição dessas amostras quanto ao uso de Português em uma realidade multilíngue. Assim, o comportamento dos fatores linguísticos – não obstante as diferenças já detalhadas – aproxima, de modo geral, o comportamento das duas variedades e, ao que tudo indica, são os fatores extralinguísticos os que atuam de maneira mais evidente sobre as variedades em questão.

Em relação a esses condicionamentos, evidenciou-se a relevância do nível de escolaridade dos indivíduos, confirmando tendências de que, quanto maior for a escolaridade, maiores serão os índices de favorecimento das marcas. Ao lado dessa condição, o emprego de línguas locais indica que a influência do menor ou maior uso do Português está, de fato, relacionado ao maior grau de instrução formal. O contato mais intenso com outras línguas, aliado à baixa escolaridade, favorece a ausência de marcas de plural nos dados observados. De outro lado, os fatores linguísticos relevantes para ambas as variedades foram a posição do sujeito e o paralelismo oracional, apontando tendências de maior favorecimento das marcas quando os sujeitos estão antepostos ao verbo e possuem marcas explícitas de número.

Por fim, afirma-se que o aprofundamento das hipóteses sobre o estabelecimento de um *continuum* de marcação de pluralidade nas variedades africanas, à semelhança do proposto em Pissurno (2017), é fundamental. Isto porque os resultados apontados nas pesquisas ora relatadas demonstram que parece mesmo haver posições intermediárias para cada uma dessas variedades, que, devido a seu caráter linguístico particular e as semelhanças em relação à aprendizagem do Português em cada sociedade, estabelecem posições que ora as aproximam e ora as afastam de seu suposto modelo de aquisição – o Português Europeu. As respostas finais para a caracterização de cada variedade dependem, sem dúvida, do maior conhecimento da complexa realidade sociolinguística das variedades são-tomense e moçambicana, o que impõe continuar nos caminhos de investigação ora trilhados.

## REFERÊNCIAS

ANTÓNIO, T. *Estratégias de ensino da concordância verbal em número à população universitária moçambicana*. 2001. Dissertação (Mestrado em Letras e Ciências Sociais) – Faculdade de Letras, Universidade Eduardo Mondlane – UEM, Moçambique.

BAVO, N. *Prática de estruturas: Concordância verbal em número*. 2011. Disponível em: <[www.catedraportugues.uem.mz/lib/docs/N.Bavo\\_FINAL20Abril2012.pdf](http://www.catedraportugues.uem.mz/lib/docs/N.Bavo_FINAL20Abril2012.pdf)>. Acesso em: 17 de dezembro de 2015.

BRANDÃO, S. F. O cancelamento da marca de número nominal na variedade urbana não standard do Português de São Tomé. *Documentos del XVI Congreso Internacional de la ALFAL*. CD-Rom. Alcalá de Henares: ALFAL, 2011.

BRANDÃO, S. F.; VIEIRA, S. R. A concordância verbal e nominal no português do Brasil e no português de São Tomé: uma abordagem sociolinguística. *Papia*, São Paulo, v. 22, n. 1, p. 7-40, 2012a.

BRANDÃO, S. F.; VIEIRA, S. R. Concordância nominal e verbal: contribuições para o debate sobre o estatuto da variação em três variedades urbanas do português. *Alfa*, São Paulo, v. 56, n. 3, p. 1035-1064, 2012b.

FIGUEIREDO, C. F. G. A concordância variável no sintagma nominal plural do Português restruturado de Almoxarife (São Tomé). *Revista de Crioulos de Base Lexical Portuguesa e Espanhola*, Lisboa, v. 1, n. 1, p. 28-55, 2009.

FIGUEIREDO, C. F. G. *A concordância plural variável no sintagma nominal do português reestruturado da comunidade de Almoxarife, São Tomé* (Desenvolvimento das regras de concordância variáveis no processo de transmissão-aquisição geracional). 2010. Tese (Doutorado em Linguística) – Faculdade de Ciências Sociais e Humanas, Universidade de Macau – Macau, 2010.

GANDRA, A. S. A concordância verbal no português europeu rural. In: OLIVEIRA, K.; SOUZA, H. C.; GOMES, L. (Org.). *Novos tons de Rosa... para Rosa Virgínia Mattos e Silva*. Salvador: EDUFBA, 2009. p. 142-161.

GONÇALVES, P. Tipologia de erros do Português Oral de Maputo: um primeiro diagnóstico. In: GONÇALVES, P.; STROUD, C. (Org.). *Panorama do Português oral de Maputo*, v. II – *A construção de um Banco de “Erros”*. INDE – Cadernos de Pesquisa n. 24. Maputo: Moçambique, 1997.

GONÇALVES, P. *Aspectos morfossintáticos da gramática do Português de Moçambique: A concordância nominal e verbal / Morphosyntactic aspects of the grammar of mozambican portuguese: Nominal and verbal agreement*. Cuadernos de la ALFAL, v. 7, p. 9-16, 2015.

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

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

JUSTINO, V. Concordância verbal em número: da descrição linguística à avaliação no contexto de ensino-aprendizagem. In: BASTOS, M.; MARQUES, J.; MONTEIRO, A.; SIOPA, C. (Org.). *Ensinar a língua portuguesa em Moçambique: Desafios, possibilida-*

des e constrangimentos. Textos selecionados das VII Jornadas da Língua Portuguesa. Porto: Porto Editora, 2015a. p. 166-186

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

LABOV, W. *Principles of linguistic change*. Oxford, Cambridge: Blackwell, 1994. v. 1.

LABOV, W. Some sociolinguistic principles. In: PAULSTON, C. B.; TUCKER, G. R. (Org.). *Sociolinguistics. The essential readings*. Oxford/New York: B. Blackwell, 2003. p. 234-250.

LUCCHESI, D.; BAXTER, A.; SILVA, Jorge A. A. da. A concordância verbal. In: LUCCHESI, D.; BAXTER, A.; RIBEIRO, I. (Org.). *O Português afro-brasileiro*. Salvador: Edufba, 2009. p. 331-371.

MONGUILHOTT, I. O. S. *Estudo sincrônico e diacrônico da concordância verbal de 3ª pessoa do plural no PB e no PE*. 2009. Tese (Doutorado em Linguística e Língua Portuguesa) – Programa de Pós-graduação em Linguística, Universidade Federal de Santa Catarina – UFSC, Florianópolis, 2009.

MONTE, A. *Concordância verbal e variação: um estudo descritivo-comparativo do Português Brasileiro e do Português Europeu*. 2012. Tese (Doutorado em Linguística e Língua Portuguesa) – Programa de Pós-graduação em Linguística e Língua Portuguesa, Universidade Estadual Paulista “Júlio Mesquita Filho” – UNESP, São José do Rio Preto, 2012.

NHONGO, N. *A concordância verbal em número no discurso escrito de estudantes universitários moçambicanos*. 2005. Tese (Licenciatura) – Faculdade de Letras, Universidade Eduardo Mondlane – UEM, Moçambique, 2005.

PETTER, M. M. T. Ampliando o *continuum* afro-brasileiro de Português. *Papia*, São Paulo, v. 25, n. 2, p. 305-317, 2015.

PISSURNO, K. C. S. *A concordância verbal de terceira pessoa do plural na variedade moçambicana do Português: uma abordagem sociolinguística*. 2017. Dissertação (Mestrado em Letras Vernáculas) – Programa de Pós-graduação em Letras Vernáculas, Universidade Federal do Rio de Janeiro – UFRJ, Rio de Janeiro, 2017.

ROMAINE, S. Introduction to the study of Bilingualism. In: ROMAINE, S. (Org.). *Bilingualism*. 2. ed. Oxford: Blackwell, 1995. p. 1-22.

RUBIO, C. F. *Padrões de concordância verbal e de alternância pronominal no português brasileiro e europeu: estudo sociolinguístico comparativo*. Tese (Doutorado em Estudos Linguísticos) – Instituto de Biociências, Letras e Ciências Exatas, Universidade Estadual Paulista “Júlio Mesquita Filho” – UNESP. São José do Rio Preto, 2012.

SCHERRE, M. M. P.; NARO, A. J. Duas dimensões do paralelismo verbal no Português popular do Brasil. *DELTA*, São Paulo, v. 9, n. 1, p. 1-14, 1993.

SCHERRE, M. M. P.; NARO, A. J. A concordância de número no Português do Brasil: um caso típico de variação inerente. In: HORA, D. (Org.). *Diversidade linguística no Brasil*. João Pessoa: Ideia, 1997. p. 93-114.

SCHERRE, M. M. P.; NARO, A. J. Mudança sem mudança: a concordância de número no português brasileiro. *SCRIPTA*, Minas Gerais, v. 9, n. 18, p. 107-129, 2006

SCHERRE, M. M. P.; NARO, A. J.; CARDOSO, C. R. O papel do tipo de verbo na concordância verbal no português brasileiro. *DELTA*, São Paulo, v. especial, n. 23, p. 238-317, 2007.

VAREJÃO, F. *Variação em estruturas de concordância verbal e em estratégias de relativização no português europeu popular*. 2006. Dissertação (Mestrado em Letras Vernáculas) – Programa de Pós-graduação em Letras Vernáculas, Universidade Federal do Rio de Janeiro – UFRJ, Rio de Janeiro, 2006.

VIEIRA, S. R. O estatuto da regra variável e o fenômeno da concordância verbal em variedades do Português. *Documentos del XVI Congreso Internacional de la ALFAL*. CD-Rom. Alcalá de Henares: ALFAL, 2011.

VIEIRA, S. R.; BAZENGA, A. Patterns of third person verbal agreement. *Journal of Portuguese Linguistics*, Lisboa, v. 12, n. 2, p. 7-50, 2013.

VIEIRA, S. R.; BAZENGA, A. A concordância da terceira pessoa plural: padrões em variedades do Português. 2013. In: VIEIRA, S. R. (Org.). *A concordância em variedades do Português: a interface Fonética-Morfossintaxe*. Rio de Janeiro: Vermelho Marinho, 2015. p. 29-75.

VIEIRA, S. R.; BRANDÃO, S. F. Tipologia das regras linguísticas e estatuto das variedades/línguas: a concordância em português. *Linguística*, Rio de Janeiro, v. 30, n. 2, p. 81-112, 2014.

VIEIRA, S. R.; SILVA, K. C. Estudo sociolinguístico da concordância verbal de 3ª pessoa do plural no Português de São Tomé e Príncipe. *Revista Letrônica*, Rio Grande do Sul, v. 10, n. 1, p. 18-29, 2017.

WEINREICH, U.; LABOV, W.; HERZOG, M. Empirical foundations for theory of linguistic change. In: LEHMANN, W.; MALKIEL, Y. (Org.). *Directions for Historical Linguistics*. Austin: University of Texas Press, 1968. p. 97-195.



# 10

CAPÍTULO

## ORDEM DOS CLÍTICOS PRONOMINAIS NO PORTUGUÊS DE SÃO TOMÉ E NO PORTUGUÊS DE MOÇAMBIQUE

SILVIA RODRIGUES VIEIRA/  
MARIA DE FÁTIMA VIEIRA

### 1. INTRODUÇÃO: DELIMITANDO OS PROPÓSITOS DESTE CAPÍTULO

O tema da colocação pronominal configura-se como importante tópico morfo-sintático para a diferenciação entre as variedades vernaculares do Português do Brasil (PB) e do Português Europeu (PE). Considerando o contexto de interações espontâneas e, portanto, supostamente sem maior controle em termos de monitoração estilística, sabe-se que colocar o pronome átono antes ou depois do verbo em início de oração ou em contextos sem a presença de algum elemento gramatical do tipo proclisador – como em *me dá um cafezinho* versus *dá-me um cafezinho* ou, ainda, em *o texto se refere* versus *o texto refere-se* – constitui, sem dúvida, expressão suficiente para a definição identitária do falante quanto ao pertencimento a uma das duas referidas variedades.

Em conjunto com o tratamento da concordância verbal e nominal (Cf. VIEIRA; BRANDÃO, 2014), objeto dos dois capítulos anteriores, descrever a colocação pronominal em variedades africanas do Português segundo os pressupostos sociolinguísticos (Cf. WEINREICH; LABOV; HERZOG, 1968; LABOV, 1972) é de fundamental relevância para a observação de propriedades linguísticas que permitam definir o estatuto de cada variedade do Português,

bem como discutir as motivações, internas ou externas à língua, para essa definição. Retomando o que já foi proposto, não só na apresentação e no prefácio deste livro, mas também nos textos precedentes, somente a descrição detalhada de dados de cada uma das variedades do Português permitirá traçar hipóteses para a configuração do chamado *continuum* afro-brasileiro (Cf. PETTER, 2007, 2015; AVELAR; GALVES, 2014).

Com o propósito de colaborar com o fornecimento de informações para tal descrição, o presente texto organiza-se da seguinte forma:

- (i) primeiramente (Seção 2), apresenta, de forma muito sintética, as tendências de colocação pronominal verificadas para o PB e o PE, sobretudo as observadas (Cf. dissertação de Mestrado de CORRÊA, 2012; tese de Doutorado de VIEIRA, 2016) em entrevistas sociolinguísticas recolhidas entre 2009 e 2011 no âmbito do Banco de dados do Projeto *Estudo comparado dos padrões de concordância em variedades africanas, brasileiras e europeias do Português*, doravante *Corpus Concordância* ([www.concordancia.lettras.ufrj.br](http://www.concordancia.lettras.ufrj.br));
- (ii) em seguida (Seção 3), reúne informações de estudos variacionistas sobre a ordem dos clíticos pronominais em duas variedades urbanas do Português na África, a de São Tomé (PST) – na amostra constituída por Tjerk Hagemeyer e disponibilizada pelo Projeto VAPOR, do Centro de Linguística da Universidade de Lisboa (cf. tese de doutorado de VIEIRA, 2016) – e a de Maputo/Moçambique (PM) – seja nas entrevistas do Projeto Panorama do Português Oral de Maputo/PPOM, em banco de dados organizado por Christopher Stroud e Perpétua Gonçalves (1997), doravante *Corpus PPOM* (Cf. tese de doutorado de VIEIRA, 2002), seja em uma pequena amostra coletada<sup>1</sup> de entrevistas sociolinguísticas realizadas recentemente, em 2016, por Silvia Rodrigues Vieira e Karen Cristina da Silva Pissurno, doravante *Corpus Moçambique*; e
- (iii) após uma breve sistematização dos resultados obtidos (Seção 4), o texto oferece reflexões para a delimitação da natureza e das propriedades das variedades do Português (Seção 5). Para tanto, leva em consideração, sobretudo, a abordagem comparativa dessas variedades, considerando não só as motivações linguísticas para a variação, mas também a situa-

---

<sup>1</sup> A coleta dos dados foi realizada pela graduanda da Faculdade de Letras da UFRJ, Amanda Carolina de Oliveira Santos, por ocasião do desenvolvimento de seu trabalho de conclusão de curso, em 2017, que versou sobre o tema da colocação pronominal.

ção de multilinguismo nas sociedades africanas em questão e na formação do Português do Brasil.

Espera-se, assim, contribuir não só com a descrição das variedades do Português, mas também com o avanço científico no sentido de explicar a complexa rede de fatores, de natureza linguística e extralinguística, que explicam a formação do perfil das variedades de uma língua.

## 2. COLOCAÇÃO PRONOMINAL NAS VARIEDADES BRASILEIRA E EUROPEIA DO PORTUGUÊS

Considerando os estudos já referidos sobre o PB e o PE com amostras contemporâneas de fala, pode-se afirmar, em linhas gerais, que a preferência na fala brasileira, quando se trata de construções com apenas uma forma verbal, é a realização pré-verbal (*me dá o cafezinho, ele se referiu ao livro*), ao contrário da europeia, que teria por preferência a realização pós-verbal (*dá-me o cafezinho, ele referiu-se ao livro*)<sup>2</sup>.

Na variedade falada no Brasil (Cf. VIEIRA, 2002; CORRÊA, 2012; VIEIRA, 2016), ocorre, na realidade, uma regra de colocação semicategórica, segundo a tipologia de Labov (2003), regra que se concretiza na quase totalidade dos dados, de 96% a 99% das ocorrências, e abrange, em termos qualitativos, amplo espectro das construções estruturais<sup>3</sup>. Nesse sentido, cabe registrar que a variante pós-verbal fica restrita no PB, em termos linguísticos, a construções particulares, como, por exemplo, as que incluem os clíticos *o, a(s)* seguidos de infinitivo (como em *encontrá-lo*) e *se* em estruturas de natureza indeterminadora (como em *trata-se de*). Ademais, em termos extralinguísticos, essas construções acabam por ser praticadas por grupos sociais também específicos, sobretudo os com maior grau de escolaridade ou acesso à cultura de escrita letrada, e em situações de interação supostamente mais monitoradas ou controladas.

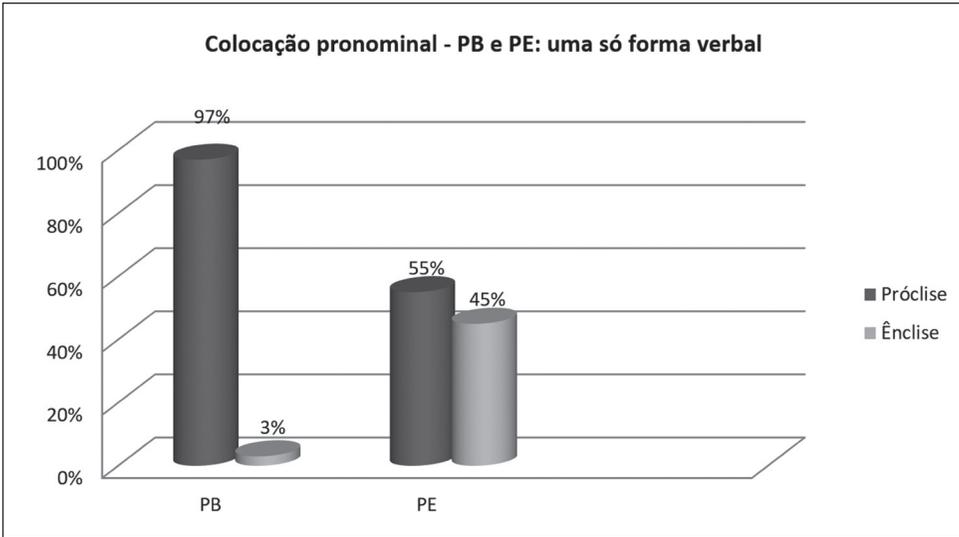
<sup>2</sup> Embora não se negue a variante mesoclítica (*dar-me-ia*), o presente artigo não a prioriza pelos seguintes motivos: (1) a mesóclise não foi registrada em diversas amostras ora em análise, o que indica sua especialização de uso, muito vinculada até onde se sabe à modalidade escrita e a gêneros textuais específicos; (2) em termos estruturais, essa variante também fica restrita a formas de futuro simples, o que também desfavorece a comparabilidade dos resultados gerais.

<sup>3</sup> VIEIRA (2014) apresenta análise detalhada do estatuto semicategórico ou variável não só da ordem dos clíticos pronominais, mas também da concordância verbal, tomando por base as variedades brasileira e europeia do Português.

VIEIRA (2016) investigou o fenômeno valendo-se de entrevistas sociolinguísticas – recolhidas entre 2009 e 2011 no âmbito do *Corpus Concordância* – realizadas com 36 informantes em cada variedade, distribuídos por sexo, três faixas etárias (18-35 anos, 36-55 anos e 56-75 anos) e três níveis de escolaridade (fundamental: 5 a 8 anos; médio: 9 a 11 anos; superior).

Em termos de distribuição geral dos dados, o estudo constata uma diferença fundamental entre as variedades: o PE apresenta ênclise categórica em início de oração, enquanto o PB registra alta produtividade de próclise (91%).

Mesmo excetuando o início absoluto de oração, as diferenças também são relevantes, como demonstram os resultados representados na Figura 1:



**Figura 1** A ordem dos clíticos pronominais em construções com uma só forma verbal na fala do PB e do PE em contexto não inicial de oração segundo VIEIRA (2016)

Fonte: Vieira (2016: 121)

Como se pode observar, os resultados confirmam as referidas tendências do PB: do total de 1.088 ocorrências de pronomes átonos em posição não inicial, a variante pré-verbal foi a opção preferencial (1.053 ocorrências, 97%); a variante pós-verbal ficou restrita a apenas 35 dados, o que corresponde a 3%. De outro lado, os dados europeus distribuem-se regularmente entre as variantes pré e pós-verbal. Do total de 2.664 ocorrências, registrou-se a variante proclítica em 1.453 dados (55%) e a enclítica em 1.211 (45%).

A autora resume, na forma de um quadro-síntese, os principais resultados a que chegou com a descrição do PB, do PE e do PST. Reproduz-se, no Quadro 1, o que foi verificado para as duas primeiras variedades:

**Quadro 1** Sistematização da ordem dos clíticos pronominais em construções com uma só forma verbal na fala do PB e do PE segundo VIEIRA (2016)

| CONSTRUÇÕES COM UMA SÓ FORMA VERBAL                  | PB  | PE  |
|--|---|---|
| Próclise em início absoluto                          | Presente e produtiva (91%)  | Ausente (0%)  |
| Distribuição das variantes nos demais contextos      | Próclise = 97%<br>Ênclise = 3%  | Próclise = 55%<br>Ênclise = 45%   |
| Estatuto da regra segundo Labov (2003)               | Semicatagórica  | Variável  |
| Variáveis extralinguísticas: condicionamento externo | Maior uso de ênclise por indivíduos com mais instrução e idade.   | Não houve relevância estatística de variáveis externas.   |
| Variáveis linguísticas                               | Registro de ênclise sobretudo com acusativos de 3ª pessoa (frequentemente combinados com infinitivos), alguns casos de <i>lhe</i> e de <i>se</i> , em estratégia de indeterminação; presença de expressões cristalizadas. | Variáveis estatisticamente relevantes: presença e natureza do elemento antecedente; tempo e modo verbal; distância entre o clítico e um elemento antecedente.   |
| Efeito proclisador                                   | Não se aplica.<br>Próclise verificada em qualquer contexto, inclusive em início absoluto.   | Efeito proclisador verificado (embora não de forma categórica): proclisadores (partículas de negação; conjunções subordinativas; preposições <i>para, de</i> ; operadores de foco ou advérbios ( <i>também, já, até</i> ) <sup>4</sup> versus não proclisadores (conjunções coordenativas, sujeitos, adjuntos adverbiais, complementos preposicionados) claramente delimitados. |

Fonte: Vieira (2016: 171-172), com adaptação

O quadro acima permite afirmar que os parâmetros de colocação do PB e do PE apresentam diferenças relevantes do ponto de vista qualitativo e quantitativo.

<sup>4</sup> Cabe enfatizar que se apresentam, aqui, tão-somente os itens verificados na amostra analisada por Vieira (2016). Para maior detalhamento da funcionalidade de elementos proclisadores no PE, recomenda-se a consulta ao capítulo “Posição dos pronomes clíticos”, de Ana Maria Martins, na *Gramática do Português* (Cf. MARTINS, 2013).

No PB, o perfil do informante que produziu os poucos dados de ênclise correlaciona-se, possivelmente, a maior idade e, secundariamente, a maior nível de instrução. No que se refere à influência de variáveis linguísticas, Vieira afirma:

os clíticos acusativos, principalmente com o verbo no infinitivo, e o pronome “se” seriam os contextos em que a posição enclítica mais ocorreria. Além disso, houve exemplos de construções cristalizadas (como “chama-se”, “deu-lhe”, “ganhou-lhe” e “parece-me”) que concretizaram a colocação enclítica. (VIEIRA, 2016, p. 173)

De outro lado, no PE, a autora observa, nos contextos não iniciais de oração, ora próclise, ora ênclise, o que, nas palavras da autora, configura uma diferença paramétrica fundamental motivada efetivamente por restrições de ordem sintática. Não tendo havido seleção de qualquer variável extralinguística, resumem-se os condicionamentos estruturais da seguinte forma:

De toda forma, é nítida a atuação diferenciada do grupo dos elementos que atuaram como proclisadores *versus* o grupo dos elementos que não atuaram como proclisadores, o que valida a interpretação de que o PE, embora em todos os casos registre variação entre as variantes pré e pós-verbal, permite a generalização de que existem contextos de (i) ênclise categórica (início absoluto de oração e de período); (ii) próclise altamente favorecida (partículas de negação, elementos subordinativos, preposições – *para, de, por* e *sem* – e operadores de foco); e (iii) próclise altamente desfavorecida (adjuntos adverbiais, complementos preposicionados, preposições – *a* e *em* –, sujeitos e conjunções coordenativas). (VIEIRA, 2016, p. 174)

No caso de construções com mais de uma forma verbal (perífrases diversas, como as temporais, aspectuais e modais), o PB admitiria como opção preferencial a posição do clítico interna ao complexo, o que corresponderia, na realidade, a uma próclise à segunda forma verbal, como fica patente em exemplos do tipo *você pode amanhã me dizer o que é preciso*. As outras posições nos complexos também ficam restritas a determinadas estruturas, como com o clítico *se* indeterminador antes da primeira (*não se pode dizer...*) ou ainda com o clítico *o, a* (s) após a segunda forma no infinitivo (*pode encontrá-lo*)<sup>5</sup>.

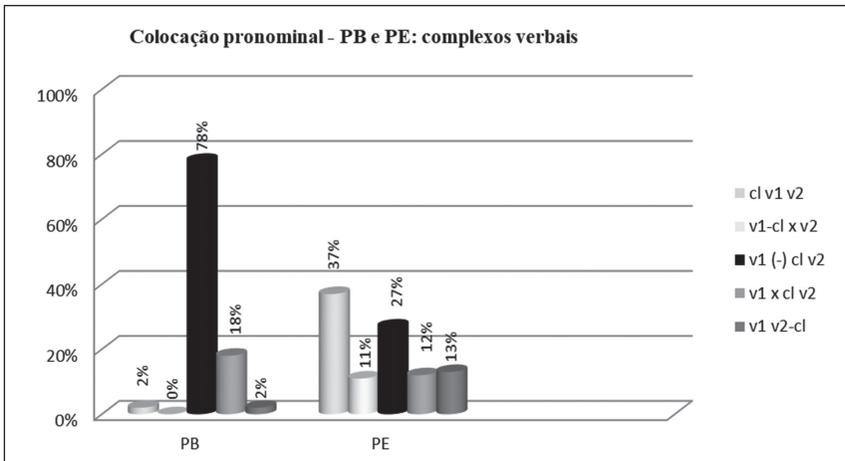
<sup>5</sup> Os resultados relativos a dados contemporâneos, de que se vale o presente artigo, confirmam, em certa medida, tendências propostas na vasta produção bibliográfica sobre o tema, cuja revisão não cabe nos propósitos limitados deste artigo. Vale destacar, aqui, que os estudos pioneiros da colocação pronominal não só apontaram as inovações brasileiras (Cf., por exemplo, PAGOTTO, 1992), mas também buscaram explicar fatores relacionados às motivações linguísticas, como as de natureza fonético-fonológica, demonstrando a correlação entre a ênclise e os

Em termos de distribuição geral dos dados de clíticos em estruturas com mais de uma forma verbal, o estudo também constata nítidas diferenças entre as variedades. A Tabela 1 e a Figura 2 permitem observar a produtividade das variantes, que controlam a posição superficial do pronome átono (cl) em relação à primeira (v1) e à segunda forma (v2) verbais, além da presença ou não de outro elemento interveniente no complexo (x).

**Tabela 1** A ordem dos clíticos pronominais em complexos verbais na fala do PB e do PE segundo VIEIRA (2016)

| COLOCAÇÃO PRONOMINAL: PB E PE |          |     |            |     |              |     |            |     |          |     |       |
|-------------------------------|----------|-----|------------|-----|--------------|-----|------------|-----|----------|-----|-------|
| Variedades                    | cl v1 v2 |     | v1-cl x v2 |     | v1 (-) cl v2 |     | v1 x cl v2 |     | v1 v2 cl |     | Total |
| PB                            | 8        | 2%  | 0          | 0%  | 289          | 78% | 66         | 18% | 6        | 2%  | 369   |
| PE                            | 246      | 37% | 68         | 11% | 172          | 27% | 75         | 12% | 81       | 13% | 642   |

Fonte: Vieira (2016: 176), adaptado



**Figura 2** A ordem dos clíticos pronominais em complexos verbais do PB e do PE segundo VIEIRA (2016)

Fonte: Vieira (2016: 176), adaptado

Da comparação dos resultados com as amostras brasileira e europeia, podem-se atestar, mais uma vez, marcantes diferenças entre as variedades. Os

pronomes com onset silábico (Cf. NUNES, 1993), e as sociais, como a vinculação da ênclise à fala monitorada de pessoas com alto nível de escolaridade (Cf. LOBO, 1992).

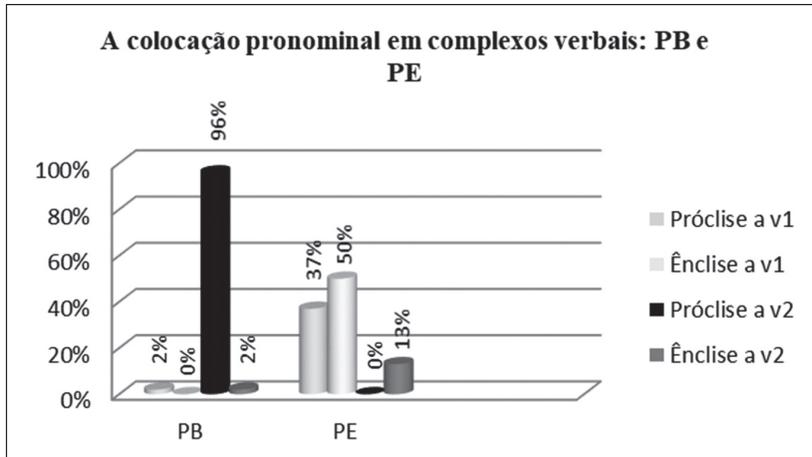
clíticos no PB distribuem-se basicamente pela posição imediatamente antecedente a v2 – em estruturas com (18%) ou sem (78%) elementos intervenientes, como, respectivamente, em *pode amanhã me dizer* e *pode me dizer* –, totalizando 96% dos casos. Poucas ocorrências de clíticos em outras posições – 2% antes de v1 (*aqui se pode dizer*) e 2% após v2 (*pode encontrá-lo*) – foram registradas. Observa-se que essas posições se restringem basicamente ao tipo de clítico; na realidade, a ligação sintática a v1 fica condicionada à estrutura indeterminadora com o pronome *se*; e a ênclise a v2 ao pronome acusativo de 3ª pessoa, manifesto na forma *lo* e variantes. Chama a atenção, ainda, a ausência de dados de clíticos antes de elemento interveniente, posição que permitiria a interpretação de ênclise a v1 (como em *começou-me a dizer*; *pode-me amanhã dizer*).

No PE, verificou-se efetiva distribuição dos dados por todas as variantes controladas, da seguinte forma: a próclise a v1 foi registrada em 37% dos dados, seguida pelo clítico imediatamente após essa forma verbal, seja em construções sem elemento interveniente (27%), seja na presença desses elementos (11%). Vale ressaltar que esta última variante, que configura caso efetivo de ênclise à primeira forma verbal do complexo, nem sequer foi registrada no PB. No PE, além da variante enclítica a v1, também foi registrada com produtividade (13%) a ênclise à segunda forma verbal. No que se refere às ocorrências da variante v1 x cl v2 (12%), Vieira (2016) enfatiza que elas ocorreram,

em geral, com os complexos verbais que já possuem algum elemento integrante em seu interior – o “que” e o “de” (*ter que/de*), por exemplo –, mostrando que, diferentemente do PB, não ocorreria uma próclise a v2, mas uma ênclise ao elemento que atua como proclisador interno ao complexo, como em “tem que se fazer”. (VIEIRA, 2016, p. 177)

Embora não se disponha de espaço nos limites do presente artigo para detalhar as construções registradas na amostra de complexos verbais, é preciso deixar claro que, nos dados da variedade europeia, houve efetivamente elementos condicionadores das variantes. A esse respeito, pode-se resumidamente afirmar que a presença de elementos proclisadores, embora favoreça a próclise ao complexo (*que me pode dizer*), não apresenta atuação categórica, tendo havido também ênclise à primeira forma verbal (em complexos com gerúndio e particípio, como em *que vem-me visitando* ou *que tinha-me encontrado*) ou à segunda (em complexos com infinitivo, como em *que pode dizer-me*). Não havendo elemento proclisador, a ênclise – a v1 ou a v2 – constitui opção natural. Clíticos imediatamente antes de v2 ficam restritos a complexos com *ter que/de* – o que não permite afirmar uma efetiva próclise a v2.

Diante do detalhamento dos resultados, a Figura 3 permite visualizar o comportamento de cada variedade em relação ao verbo que está efetivamente hospedando o clítico em cada caso.



**Figura 3** A ordem dos clíticos pronominais em complexos verbais do PB e do PE (variantes amalgamadas) segundo VIEIRA (2016)

Fonte: Vieira (2016: 176), adaptado

Como se pode observar, a regra de colocação nos complexos verbais do PB, assim como se atestou em construções com uma única forma verbal, não se apresenta efetivamente variável, mas semicategórica, registrando dados que podem ser interpretados como preferencialmente próclíticos a v2. A ligação sintática do clítico a v1 e a ênclise a v2 ficam condicionadas basicamente ao tipo de pronome. De outro lado, os dados de complexos verbais no PE apresentam colocação pronominal sensível às variáveis estruturais referentes ao contexto morfossintático em questão.

### 3. A COLOCAÇÃO PRONOMINAL EM VARIEDADES URBANAS AFRICANAS: PST E PM

Cumprindo o objetivo de apresentar descrições de variedades urbanas do Português, esta seção exhibe resultados da ordem dos clíticos pronominais com base em amostras do Português falado em variedades africanas, quais sejam: a variedade de São Tomé (PST) e a de Maputo/Moçambique (PM). Trata-se de variedades em que igualmente o Português coexiste com várias outras línguas, embora mediante distintas realidades sociolinguísticas, conforme atesta Hagemeyer, no prefácio da presente obra: a moçambicana “está em contacto com línguas aglutinantes do grupo banto e apresenta uma taxa de nativização crescente, embora ainda relativamente baixa”, enquanto a são-tomense “constitui a L1 da

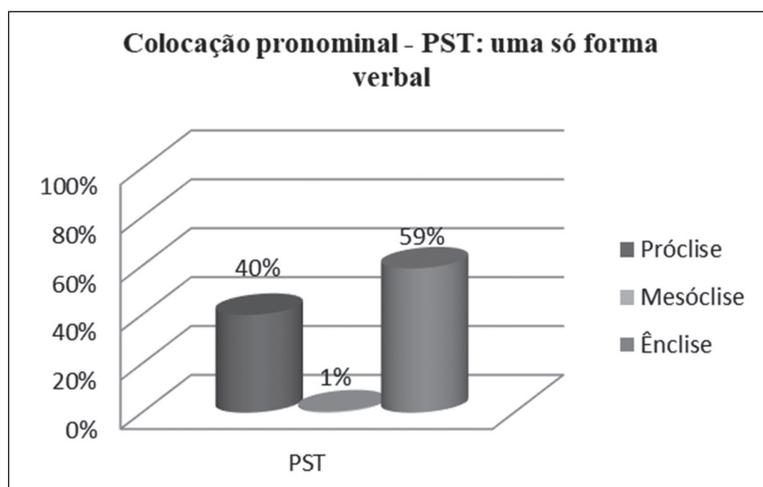
maioria da população, mesmo que os censos nacionais não explicitem esta informação”<sup>6</sup>.

### 3.1 A colocação pronominal na variedade urbana de São Tomé

Embora não se tenha por propósito, nos limites deste artigo, proceder a uma acurada revisão da literatura sobre o tema em cada variedade analisada, é notória a limitada produção bibliográfica sobre a ordem dos clíticos pronominais no PST<sup>7</sup>. Esta seção baseia-se, portanto, conforme já anunciado, em resultados de análise do *Corpus VAPOR* também empreendida por Vieira (2016).

#### 3.1.1 Construções com uma só forma verbal

Considerando as construções com uma única forma verbal, a Figura 4 permite observar a distribuição dos 525 dados em análise.



**Figura 4** A ordem dos clíticos pronominais em complexos verbais na fala do PST segundo VIEIRA (2016)

Fonte: Vieira (2016: 121)

<sup>6</sup> A respeito do multilinguismo em São Tomé e Príncipe, vale consultar Hagemeyer (2009); quanto ao estatuto da variedade moçambicana do Português, também constitui leitura indispensável Gonçalves (2010).

<sup>7</sup> Na revisão da literatura feita por Vieira (2016), faz-se alusão apenas à dissertação de Mestrado de Rita Gonçalves (2009), intitulada *A Colocação dos Pronomes Clíticos no Português Oral de S. Tomé: Análise e discussão de contextos numa perspectiva comparativa*, também desenvolvida com dados do Projeto VAPOR, já mencionado na introdução do presente texto.

Como se pode observar, houve um número um pouco maior de dados de ênclise (310/525; 59%) do que de próclise (215/525; 40%) na amostra são-tomense. A distribuição dos dados alinha, portanto, o perfil dessa variedade mais ao do PE do que ao do PB, no sentido de que o fenômeno se revela variável e não semi-categórico, conforme se apresentou na seção anterior. Os exemplos 1 e 2 permitem verificar a referida variação:

Ex.1: fui trabalhar na casa de pessoa porque salário era muito baixo, também com a ajuda de minha avó sempre ela apoiou-me [PSTA1M]<sup>8</sup>

Ex.2: eu tento lutar também porque a gente não pode sentar só esperar o que Deus vai dar né Deus, bíblia diz, Deus disse põe tua mão que eu te ajudo então eu ponho mão [PSTA1M]

Diferentemente de outras variedades controladas no estudo de Vieira (2016), encontrou-se uma ocorrência de mesóclise (1/525; 1%) na amostra de São Tomé:

Ex.3: MAS é preciso que haja motivação por parte dos governantes e essa motivação tornar-se-ia mais extensiva aos professores [PSTC3H]

Em termos de comparação geral entre as variedades brasileira, europeia e são-tomense, Vieira afirma que:

O PST, de acordo com a análise, segue tendências gerais equivalentes às verificadas para o PE, quais sejam: (i) o fenômeno também se mostra variável; (ii) ocorre ênclise categórica em início absoluto de oração e de período (diferenciando-se ambas do PB radicalmente nesse contexto) e (iii) apresenta algumas variáveis linguísticas como condicionamentos favorecedores da próclise consoante as mesmas tendências gerais. (VIEIRA, 2016, p. 174)

Não obstante essa esperada semelhança geral entre as variedades europeia e são-tomense – visto ser a primeira, em tese, a norma de referência da segunda –, a autora, ao detalhar o comportamento dos dados quantitativa e qualitativamente, aponta importantes particularidades do PST.

Primeiramente, embora as variáveis linguísticas selecionadas – com exceção de *Tonicidade da forma verbal* – sejam as mesmas atuantes nos dados da amostra europeia no trabalho de Vieira (2016) – *Elemento proclisador; Tempo/modo verbal; Distância entre o clítico e o termo antecedente* –, a observação detalhada das

<sup>8</sup> Para a identificação do perfil dos informantes, a notação que se faz ao fim de cada exemplo corresponde ao perfil quanto a variedade (PST = São Tomé; PM = Moçambique/Maputo); idade (A = 18 a 35 anos; B = 36 a 55 anos; C = de 56 anos em diante); escolaridade (1 – até 8 anos de escolaridade; 2 – cerca de 11 anos de escolaridade; 3 = curso superior completo); sexo (H = homem; M = mulher). No caso da variedade moçambicana, identifica-se, ainda, a amostra em questão.

variáveis e de cada contexto morfossintático particular permitiu aferir o comportamento típico do PST. Em segundo lugar, chama atenção a relevância estatística de variáveis extralinguísticas, quais sejam: *Sexo*, selecionada em primeiro lugar, e o grupo intitulado *Frequência de uso de um crioulo*, controlada segundo o depoimento dos próprios informantes.

O Quadro 2 apresenta, de forma resumida, os principais condicionamentos de cada variável linguística controlada.

**Quadro 2** Sistematização da ordem dos clíticos pronominais em construções com uma só forma verbal no PST segundo VIEIRA (2016): variáveis linguísticas

| CONSTRUÇÕES COM UMA SÓ FORMA VERBAL                | FATORES FAVORECEDORES DA PRÔCLISE   | FATORES DESFAVORECEDORES DA PRÔCLISE  |
|--|---|---|
| Elemento antecedente                               | Partículas de negação<br>Elementos subordinativos<br>Preposições<br>Operadores de foco (com pouca expressividade) | Conjunções coordenativas<br>Sujeitos<br>Adjuntos adverbiais<br>Compl. preposicionados |
| Tempo/modo verbal                                  | Formas/tempos do subjuntivo   | Formas/tempos do indicativo<br>Infinitivo<br>Imperativo                               |
| Distância entre o clítico e o elemento proclisador | Ausência de elementos intervenientes  | Presença de uma ou mais sílabas intervenientes  |
| Tonicidade da forma verbal                         | Oxítonos  | Paroxítonos e proparoxítonos  |

Fonte: Vieira (2016: 171-172), adaptado

A análise demonstra que o PST tende a utilizar, de modo geral, mais a variante enclítica do que a proclítica, sendo o início de oração contexto de realização categórica, assim como ocorre no PE. No que se refere aos elementos proclisadores, verifica-se que a atuação dos fatores também aproxima PE e PST: com exceção dos operadores de foco, que não se manifestaram como fator de próclise na amostra são-tomense de forma tão evidente, os elementos atuantes são basicamente os mesmos.

Observando, entretanto, os índices matemáticos providos pela análise estatística (consoante o pacote de programas Goldvarb-X), Vieira (2016) relata que a ação dos elementos proclisadores não revela “uma nítida dicotomia entre elementos proclisadores *versus* elementos não proclisadores, como ocorre no PE (com certa variação), mas há forte oscilação na colocação nos mesmos contex-

tos sintáticos” (p. 217). Em outras palavras, as diferenças entre elementos favorecedores e desfavorecedores de próclise não seriam tão nítidas, exibindo distâncias menores. A autora declara que, no mesmo ambiente morfossintático, tanto a próclise como a ênclise são naturalmente concretizadas, como se pode verificar nos exemplos 4 a 9:

Ex. 4: por exemplo, se formos para médico, eu já fui para para para para hospital hhh naquilo que chama-se de urgência [PSTA2H]

Ex. 5: posso apoiar também qualquer amigo qualquer amigo que me peça, vamos lá [PSTA1H]

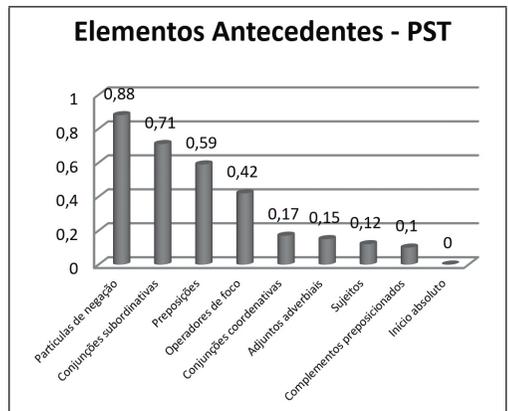
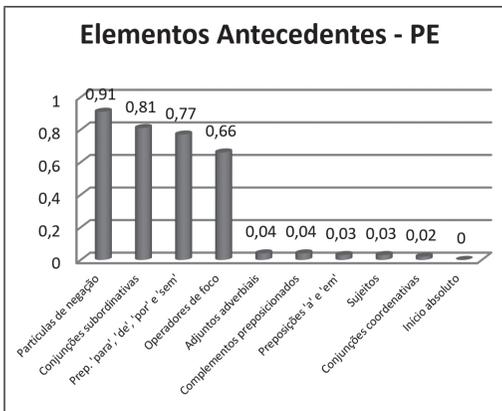
Ex. 6: em termos de política (como que) oposição aí tinha problemas era mais no sentido/ ( ) não era porque o povo tava oprimido que não se podia nada não falava-se... fazia-se tudo (normal) [PSTC2H]

Ex.7: A: Não, lá não, não se paga nada. hhh todo o mundo é bem-vindo [PSTA1H]

Ex. 8: dos dois um de química: e um de biologia ((risos)) Eu lembro-me é [PSTB3M]

Ex. 9: foi cubano... o de biologia no décimo primeiro também foi cubano... eu me lembro do/ dos dois professores que tive [PSTB3M]

Essa atuação menos rigorosa dos elementos proclisadores no PST em relação ao que ocorre no PE pode ser observada nas Figuras 5 e 6:



**Figuras 5 e 6** Atuação da variável *Elementos antecedentes* no PE e no PST segundo Vieira (2016)

Fonte: Vieira (2016: 162)

Como se pode observar, enquanto o efeito proclisador nos dados do PE pode ser visualmente delimitado em dois grupos, ele ocorre no PST de forma mais gra-

dual: a imagem deixa ver o decréscimo paulatino<sup>9</sup> da atuação dos chamados fatores de próclise.

Além desse comportamento em certa medida instável, verifica-se, também, no PST a presença da variante pré-verbal diante de elementos que não atuam, até onde se sabe, como proclisadores no PE, como, por exemplo, conjunções coordenativas (exemplo 10) e SN sujeito (exemplo 11).

Ex. 10: lá naquelas tascas, não os altos restaurantes mas lá onde os pobres comem, viver e sentir o calor da cozinha, ver as panelas, enfim, cheias de tisma de de forno e de carvão mas me senti como se tivesse sido em casa [PSTA3H]

Ex. 11: Eu consegui conviver com a população, consegui mergulhar onde há pobreza em Nigéria, as pessoas me olhavam com bons olhos, consegui comer com os nigerianos [PSTA3H]

Exemplos como esses, considerados usualmente na literatura como inovações brasileiras – em que se verifica a próclise em contextos sem tradicionais proclisadores –, aliados a essa certa oscilação no comportamento quanto à ordem aproximam curiosamente certos dados do PST aos do PB. Gonçalves (2009), em estudo do tema com dados do PST, propõe que essa instabilidade é típica “de uma língua que ainda não está completamente formada, em especial em um contexto de plurilinguismo como o que existe em São Tomé”, ou, ainda, “que, embora actualmente comece a afirmar-se como língua materna, foi adquirida como língua segunda” (GONÇALVES, 2009, p. 22). Por essa interpretação, a pesquisadora propõe que o Português em São Tomé “constitui uma interlíngua ainda não fossilizada pelo que a variação nas mesmas estruturas não nos permite caracterizar o sistema de colocação dos clíticos mas tão somente descrevê-lo o mais pormenorizadamente possível.” (GONÇALVES, 2009, p. 22).

Quanto às demais variáveis selecionadas, as tendências atestadas estão perfeitamente de acordo com as hipóteses formuladas: formas do subjuntivo, que ocorrem comumente na presença de elementos proclisadores, favorecem a variante pré-verbal, enquanto as demais a desfavorecem; elementos proclisadores atuam mais fortemente quando próximos ao verbo.

No que se refere especificamente à tonicidade das formas verbais, segundo o qual as formas oxítonas favorecem a próclise e as paroxítonas ou proparoxítonas a desfavorecem, Vieira associa esse resultado à hipótese de “a ligação fonológica do clítico no PST também ser para a esquerda, como supostamente ocorre com o

<sup>9</sup> Vieira (2016) demonstra essa diferença entre as duas variedades pelo cálculo do range, medida que avalia as distâncias entre os índices relativos de cada fator.

PE, formando, assim, vocábulos fonológicos com o que vem antes (ênclise fonológica) e não depois (próclise fonológica), como ocorre no PB.” (VIEIRA, 2016, p. 169). Nesse sentido, estaria reforçada a tendência à ênclise em geral, o que acaba por formar proparoxítonas (*mandaram-me*) e até esdrúxulas (*ajudávamos*), como as registradas nos exemplos 12 e 13.

Ex. 12: porque o ministério conhece sabe que eu eu tenho uma segunda profissão que eu pratico de vez em quando eu acho que foi nessa base que mandaram-me pra aqui [PSTB3H]

Ex. 13: também naquela altura eu lembro-me do velho Costa Alegre o ( ) Aguiar e: muitos outros muitos outros (enfim) ajudávamos muito no conhecimento da língua [PSTC3H]

Ainda que apontando a falta de dados e de evidências empíricas para qualquer conclusão de ordem rítmica e defendendo maior aprofundamento quanto a esse tipo de motivação, Vieira (2016) aventa a seguinte hipótese quanto aos resultados da variável *Tonicidade da forma verbal*:

Ademais, isso pode justificar o fato de não haver clíticos em início absoluto de oração e de período, pois a ligação fonológica do clítico no PST seria para a esquerda – como no PE –, ao contrário do PB, em que a ligação fonológica seria para a direita, como mostram os trabalhos de Vieira (2002) e Corrêa (2012). (VIEIRA, 2016, p. 169)

Quanto ao condicionamento extralinguístico, chama atenção a relevância estatística da variável *Sexo do informante*, cujos resultados se expõem na Tabela 2:

**Tabela 2** A ordem dos clíticos pronominais de acordo com o *Sexo do informante* em construções com uma só forma verbal no PST segundo Vieira (2016)

| SEXO DO INFORMANTE | VALOR ABSOLUTO | PERCENTUAL | PESO RELATIVO |
|--------------------|----------------|------------|---------------|
| Homem              | 137/260        | 53%        | .57           |
| Mulher             | 78/265         | 29%        | .40           |

Fonte: Vieira (2016: 155)

Os resultados demonstram que a fala indivíduos do sexo masculino favorece a variante próclítica (.57), em comparação à do sexo feminino (.40). Essa diferença de comportamento só pode ser seguramente interpretada com base em informações mais precisas com relação a aspectos tanto linguísticos quanto extralinguísticos.

Primeiramente, é preciso detalhar os contextos de utilização de próclise e de ênclise na subamostra de fala feminina e na masculina. A fim de detalhar esses

contextos, Vieira (2016) procede ao cruzamento entre os grupos de fatores *Sexo* e *Natureza do elemento antecedente*. Consoante os resultados obtidos, foi possível verificar que a diferença mais expressiva de comportamento entre homens e mulheres está relacionada ao emprego de próclise ou ênclise diante dos elementos que não são tradicionalmente considerados proclisadores. Ao que parece, os homens, na busca de concretizar o efeito proclisador, acabam por generalizar a tendência e realizar mais a variante pré-verbal mesmo em contextos sem efetivos elementos proclisadores. Isso pôde ser verificado no caso das conjunções coordenativas (H – 43% / M – 9%), adjuntos adverbiais (H – 67% / M – 6%), sujeitos (H – 25% / M – 14%) e complementos preposicionados (H – 25% / M – 0%), diante dos quais os informantes do sexo masculino (H) realizaram mais próclises do que as mulheres (M).

Em termos sociais, entende-se ser fundamental interpretar os resultados da variável *Sexo* mediante o conhecimento relativo à inserção social de mulheres e homens na sociedade são-tomense, conhecimento de que não se dispõe efetivamente neste trabalho. A diferença verificada, segundo Vieira (2016, p. 156), “poderia estar vinculada ao fato de as mulheres, para se enquadrarem melhor na sociedade, utilizarem, ao máximo possível, as regras presentes na gramática do PE.”. Consoante essa hipótese, haveria relações distintas entre homens e mulheres em relação ao modelo europeu de colocação justificadas pela busca de prestígio no domínio da Língua Portuguesa. Certamente, a aferição dessa hipótese necessita de aprofundamento no conhecimento das relações na sociedade são-tomense.

Para o conhecimento da apropriação social da ordem dos clíticos na sociedade são-tomense, os resultados obtidos para a variável *Frequência de uso de um crioulo* podem ser esclarecedores. Essa variável investiga se os padrões de uso da colocação pronominal no PST estão relacionados à relação do usuário com o emprego, mais ou menos frequente, do Português ou da língua local, o Forro. Com base em Brandão (2011), os dados foram codificados de modo a controlar os seguintes fatores: (i) frequência zero ou baixa de uso do crioulo, quando os indivíduos declaram que se expressam fundamentalmente em Português; (ii) frequência média, quando os indivíduos declaram que se expressam em Português, mas dominam um crioulo e dele fazem uso eventualmente; e (iii) frequência alta, quando os indivíduos, embora falem o Português, declaram que se expressam, regularmente, num crioulo.

Os resultados obtidos não permitiram confirmar cabalmente a hipótese de que, quanto maior o uso do Português, maior o domínio da norma europeia de referência. Antes, sugerem que uma pessoa que usa frequentemente o crioulo seria a que mais produziria a variante proclítica (0,89) – o que pode sinalizar uma tendência a generalizar a próclise por desconhecimento do efeito proclisador; em

segundo lugar no que se refere ao uso da próclise, viriam os que supostamente não utilizariam o crioulo (0,58) e, por último, os falantes que só o utilizam às vezes, mostrando que a próclise seria desfavorecida nesse caso (0,39).

A análise desses resultados ainda não pode ser seguramente proposta em virtude dos limites metodológicos da pesquisa. Conforme propõe Vieira (2016),

efetivamente, qualquer interpretação das interinfluências do uso do Forro dependeria de pesquisa específica com amostra que controlasse efetivamente esse efeito, além de pesquisa etnográfica a respeito dos informantes e de sua relação com o uso da Língua Portuguesa e da língua local (p. 171).

De todo modo, a correlação entre esse grupo de fatores e a colocação pronominal mostrou-se relevante, o que permite traçar hipóteses para etapas futuras investigativas. Nesse sentido, Vieira (2016) observa:

Retirando o único informante que se declara como um falante do crioulo com alta frequência, verifica-se que os que afirmam utilizar basicamente o Português teriam uma preferência pela variante proclítica (possivelmente por dominar os contextos europeus de próclise), e os que conhecem e utilizam o Forro, mesmo que eventualmente, prefeririam a colocação enclítica, o que talvez possa sinalizar a influência da colocação pós-verbal dos pronomes objetos do crioulo, segundo Ferraz (1979) e Hagemeyer (2007). (VIEIRA, 2016, p. 170-171)

### 3.1.2 Construções com mais de uma forma verbal

A análise do tema no contexto de complexos verbais, que contou com um conjunto de 120 dados (cf. Tabela 3; Figura 7), mostra a produtividade das variantes, conforme a posição superficial e a presença ou não de material interveniente, na variedade são-tomense.

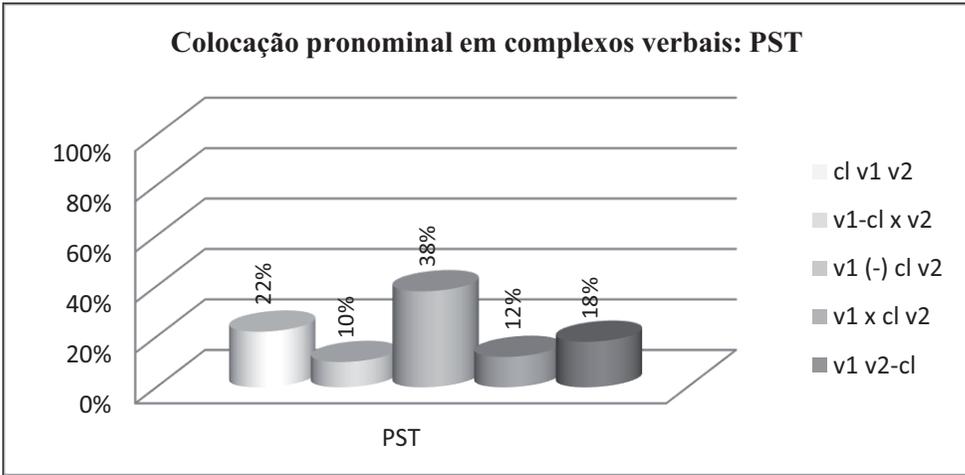
**Tabela 3** A ordem dos clíticos pronominais nos complexos verbais do PST segundo Vieira (2016)

| A ORDEM DOS CLÍTICOS NOS COMPLEXOS VERBAIS PELA FORMA DE V2: PST |          |     |            |     |             |     |            |     |          |     |       |
|--|----------|-----|------------|-----|-------------|-----|------------|-----|----------|-----|-------|
| Variedade  | cl v1 v2 |     | v1-cl x v2 |     | v1 (-)cl v2 |     | v1 x cl v2 |     | v1 v2 cl |     | Total |
| PST  | 26       | 22% | 12         | 10% | 45          | 38% | 14         | 12% | 23       | 18% | 120   |

Fonte: Vieira (2016: 176)

A distribuição dos dados revela que, no PST, assim como no PE, ocorrem realizações de todas as variantes controladas, tendo sido a v1-cl v2 a mais produ-

tiva (38%), seguida pela próclise a v1 (22%) e pela ênclise a v2 (18%). Na presença de elemento interveniente (x), ocorreram as colocações adjacentes a v1 (10%) e a v2 (12%). Nesse caso, é importante verificar o contexto em que as variantes se realizam, de modo a saber se ocorreria, por hipótese, a chamada próclise a v2, como no PB, ou se haveria apenas uma construção específica, como ocorre no PE, em que esse uso fica restrito a construções do tipo *ter quede* mais infinitivo.



**Figura 7** A ordem dos clíticos pronominais nos complexos verbais do PST segundo Vieira (2016)

Fonte: Vieira (2016: 176)

Primeiramente, é necessário observar que a maioria das ocorrências na amostra são-tomense é constituída de complexos com v2 infinitivo, tendo sido registrados apenas 11 dados de complexos com gerúndio e 5 com participío, conforme se pode verificar na Tabela 4.

**Tabela 4** A ordem dos clíticos pronominais nos complexos verbais de acordo com a forma do verbo principal no PST segundo Vieira (2016)

| A ORDEM DOS CLÍTICOS NOS COMPLEXOS VERBAIS PELA FORMA DE V2: PST |          |     |            |     |              |     |            |     |          |     |       |      |
|--|----------|-----|------------|-----|--------------|-----|------------|-----|----------|-----|-------|------|
| V2   | cl v1 v2 |     | v1-cl x v2 |     | v1 (-) cl v2 |     | v1 x cl v2 |     | v1 v2 cl |     | Total |      |
| Infinitivo   | 22       | 21% | 12         | 11% | 33           | 33% | 14         | 13% | 23       | 22% | 104   | 87%  |
| Gerúndio   | 2        | 18% | 0          | 0%  | 9            | 82% | 0          | 0%  | 0        | 0%  | 11    | 9%   |
| Participío   | 2        | 40% | 0          | 0%  | 3            | 60% | 0          | 0%  | 0        | 0%  | 5     | 4%   |
| Total  | 26       | 22% | 12         | 10% | 45           | 38% | 14         | 12% | 23       | 18% | 120   | 100% |

Fonte: Vieira (2016, p. 203)

Verifica-se que a opção preferencial no caso das construções com gerúndio e particípio é efetivamente a posição do clítico interna ao complexo, cuja ligação para v1 ou para v2 não se pode determinar *a priori*.

Ex. 14: uma vez que fazem aquilo que entendem não trabalham e se não trabalham não há produção e assim a pobreza vai se agravando portanto [PSTB3H]

Ex. 15: G: Exactamente, tem se feito algum esforço através da música [PSTA3H]

Tanto em complexos com gerúndio quanto em complexos com particípio, foram registradas, ainda, duas ocorrências de variante proclítica a v1 (como nos exemplos 16 e 17).

Ex. 16: nós temos que: tentar fazer de conta eh olhando um pouco também pra: pra transformações que se foram dando que se foram dando que é o caso perca de valores não é? [PSTB3M]

Ex. 17: a vida tem sempre altos e baixos momentos, bons e maus momentos, e falando sobre reportagem eu não estou a ver aqui muitas que me tenham marcado de forma chocante [PSTA3H]

Exemplos como esses – produzidos por indivíduos com curso superior – sugerem que o padrão de colocação pronominal são-tomense em complexos verbais seguiria as tendências do PE, de modo a efetivar o efeito proclisador também em construções com mais de uma forma verbal. Entretanto, a amostra, embora pequena, também apresentou dados (como nos exemplos 18, 19, 20) que contrariam essa tendência.

Ex. 18: eu acho que tem entrado muitos nigerianos e que não há controlo, eu não acredito que tem se feito uma avaliação [PSTB2M]

Ex. 19: É verdade que em tempos atrás, no período... no período, antes, né hhh, eu acho que já estou me perdendo um pouco [PSTA3H]

Ex. 20: a gente ouve um monte de besteira porque não estamos, não está a entender o que é que está se passando, cada um fala aquilo que ele imagina [PSTA1M]

Observar o conjunto de dados com complexos com v2 infinitivo (104 ocorrências) permitirá interpretar melhor a atuação das partículas proclisadoras nos dados de PST. Nessas construções, todas as variantes foram registradas.

Os 12 dados da ordem v1-cl x v2, com a presença de elementos intervenientes, constituem construções de natureza aspectual do tipo *estar a + infinitivo*, *ir a + infinitivo*, *por a + infinitivo*, *habituar + infinitivo*, *começar a + infinitivo* e, ainda, a estrutura modal *ter que + infinitivo*, todas elas combinadas com o pronome *se*, como nos exemplos 21 e 22.

Ex. 21: o paludismo, o mosquito causador de paludismo quase que não se vê. Depois de terminar as pulverizações, dois, três, quatro meses começa-se a sentir novamente o ressurgir dos mosquitos [PSTA3M]

Ex. 22: A: nem por isso. Nem por isso. Está-se a complicar cada dia [PSTA3M]

Dos 14 dados com a variante v1 x cl v2, dez são do tipo *ter que + infinitivo*, conforme se atestou nos dados do PE. Chamam a atenção, entretanto, 4 ocorrências que não são desse tipo, apresentadas a seguir.

Ex. 23: não sei o que é que estava por detrás das coisas mas, infelizmente, essas pessoas lá não estavam a se revelar muito simpáticas [PSTA3H]

Ex. 24: não é que eu não tenha mais para dar, mas eu acho que eu dei o meu contributo, eu dei o meu contributo, eu estou a me licenciar agora para poder aplicar [PSTA3H]

Ex. 25: depois estava a dar entrevista ontem a dizer não sei quê não sei quê que é culpa de não sei quê. É culpa de quem? Ele é que é culpado. O indivíduo pode até se exprimir assim, abertamente [PSTA3H]

Ex. 26: Os mais velhos, esses infelizmente estão a morrer, né, eles que têm mais conhecimento disso, estão a morrer e não estão a... os jovens hoje em dia não estão a se interessar que é para aprender essa língua [PSTA3M]

Ao que tudo indica, a tendência mais geral do PST é a de registrar a ênclise a V1, assim como se verificou no PE. Os exemplos ora, em análise, trazem, no entanto, uma reflexão adicional acerca da oscilação de comportamento dos dados do PST. Os exemplos que contêm a preposição *a* interveniente particularizam o comportamento da variedade são-tomense em relação à europeia. Segundo Gonçalves (2010, p. 52-53), procedendo à listagem das características diferentes do Português em Moçambique, essas estruturas seriam traços distintivos de PM e PE, visto que seria típico da variedade europeia a posição do clítico após v2 (*estavam a revelar-se; estou a licenciar-me; estão a interessar-se*). No exemplo 25, em especial, chama atenção a adjacência do pronome a v2 após o elemento interveniente *até*. Essa construção costuma ser apontada, costumeiramente, como a inovadora próclise a v2, registrada no PB.

Conforme destaca Vieira (2016),

(...) é digno de nota o registro dessa ocorrência na amostra são-tomense, em termos qualitativos, mesmo sendo apenas um dado. Em última análise, essa ocorrência pode sugerir que os informantes são-tomenses admitam, ainda que raramente, o pronome proclítico a v2, a variante considerada uma inovação brasileira, sendo, assim, uma particularidade do PST, não verificada

na amostra do presente trabalho em relação ao PE (nem em trabalhos anteriores como os de Vieira (2002) e Vieira, M. F. (2011)). Caso realmente se confirme ser esse dado natural no PST, aumenta a confiabilidade da hipótese de que a instabilidade advinda da situação de multilinguismo que envolve a implementação da Língua Portuguesa em São Tomé, por causa da intensa situação de contato linguístico, possa explicar as feições diferenciadas de uma língua transplantada em relação à língua do colonizador, como também aconteceu no caso do PB. (VIEIRA, 2016, p. 204-205)

No que se refere à posição do clítico antes do complexo verbal, em 22 dados com o infinitivo, cabe registrar que ela sempre ocorreu após a presença de elemento proclisador. No entanto, assim como no PE, o efeito proclisador não se deu de forma categórica, de modo que, mesmo estando o clítico antecedido de elementos subordinativos, partículas de negação e operadores de foco, foram registradas as variantes enclítica a v1 ou a v2, como se pode verificar nos exemplos 27 a 29.

Ex. 27: Eu se estiver a precisar de dois mil dólares para, pronto, só um exemplo, para editar um disco, o mínimo que **possam me dar** [PSTA3H]

Ex. 28: A: Eles não **vão me perseguir** porque eu não tenho dinheiro. [PSTB1M]

Ex. 29: sinto bem é na minha casa origem... é: eu fico fico mesmo antes que **venha me fazer** mais confusão a cabeça [PSTB3M]

A colocação depois do complexo verbal (como nos exemplos 30 e 31), em 23 dados com o infinitivo, ocorreu com diversos tipos de complexos verbais e pronomes (*me, se, o(s) e lhe*), em estruturas com elementos proclisadores ou não, como ocorreu no PE.

Ex. 30: eu julgo que... para pessoa... do do meu cariz, a pessoa do meu próprio tipo, eu acho que **devo voltar-me** mais para a sociedade [PSTA3H]

Ex. 31: se não te: m é claro que não / (ela) não **pode manifestar-se** né? [PSTB3M]

Em todos os tipos de complexos, foram encontrados dados em que há um elemento proclisador e que não ocorre a variante anterior a v1. Dados desse tipo confirmam o que se estabelece em Martins (2013) para o PE: no caso de complexos verbais, o pronome átono pode ocorrer, quando há um elemento proclisador, não só antes do complexo, mas também depois dele. No PST, essa variação parece ser ainda mais intensa, visto que a oscilação do efeito proclisador, que foi verificada nas estruturas com uma só forma verbal, pode ter sido intensificada nos complexos verbais. Ademais, chama a atenção, ainda que em poucos dados, o

registro de clíticos em posições até então não verificadas no PE, como a suposta próclise a v2.

De modo geral, constata-se que, enquanto não se verifica na amostra referente ao PB efetiva variação e a regra é do tipo semicategórico (LABOV, 2003), dado seu perfil quantitativo e qualitativo, o comportamento dos dados do PE e do PST se correlaciona ao de uma regra variável, com comportamentos particulares a depender da forma de v2.

### 3.2 A colocação pronominal na variedade urbana de Moçambique

Reportam-se, nesta seção, resultados de investigações com entrevistas sociolinguísticas realizadas em Maputo, capital do país, recortando-se, para o presente artigo, exclusivamente os resultados referentes às construções com uma só forma verbal<sup>10</sup>. Apresentam-se, primeiramente, os resultados da tese de Doutorado de Vieira (2002) – que observou o fenômeno no *Corpus PPOM*, com entrevistas realizadas nos anos de 1990 – e, em seguida, as tendências gerais verificadas em uma amostra reduzida (com apenas seis entrevistas) do *Corpus Moçambique*, constituído em 2016.

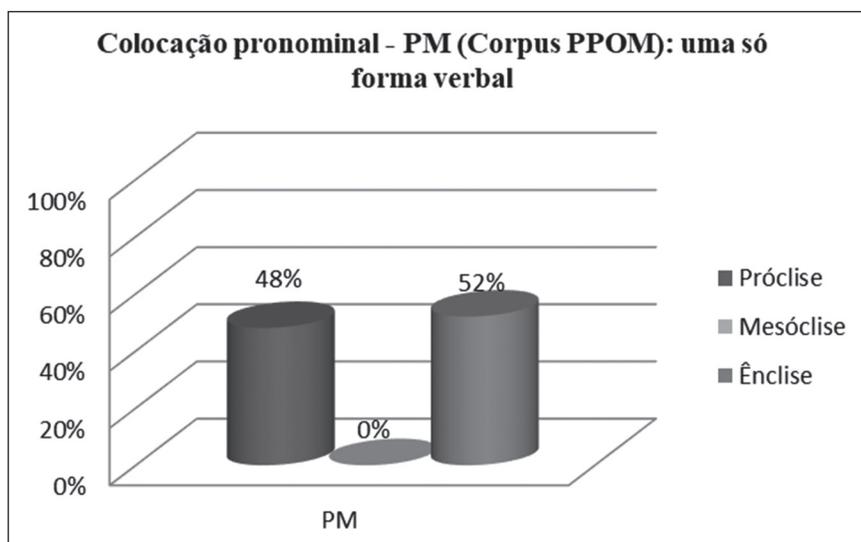
Vieira (2002) analisa, do conjunto de entrevistas do *Corpus PPOM*, 1347 ocorrências da ordem dos clíticos pronominais na fala de 38 entrevistados, distribuídos, quanto à faixa etária, em três grupos – faixa A (18 a 35 anos), faixa B (36 a 55 anos), faixa C (56 anos em diante). Dadas as particularidades no processo de aprendizagem do Português, a autora respeitou os níveis de escolaridade propostos pelos organizadores do referido *corpus*: nível 1 – da 3ª a 7ª classe –; nível 2 – da 8ª a 11ª classe; nível 3 – frequência universitária. Na medida do possível, foram utilizadas cinco entrevistas por faixa etária e nível escolaridade.<sup>11</sup>

A distribuição das 1347 ocorrências pelas variantes pré e pós-verbal ocorreu conforme se detalha na Figura 8.

---

<sup>10</sup> Não se apresentam os resultados para os complexos verbais considerando o *Corpus Moçambique*, primeiramente porque a análise da ordem dos clíticos nessas construções ainda não foi concluída. Ademais, os resultados apresentados por Vieira (2002), com base na análise de dados do *Corpus PPOM*, foram adquiridos com base em metodologia – a partir de regra binária de colocação – diferente da adotada por Vieira (2016) – a partir de regra enéaria – para o PST, o que não facilitaria a comparação dos resultados no presente artigo.

<sup>11</sup> Não foi possível, no estudo de Vieira (2002), valer-se de cinco entrevistas por faixa etária no caso do grau de instrução máximo (frequência universitária), o que se justifica pelas limitações do próprio *Corpus PPOM*, que não dispunha de mais inquéritos do referido perfil além dos utilizados. Também não foi possível selecionar os informantes segundo a variável *Sexo do informante*, visto que os inquéritos não foram distribuídos uniformemente entre homens e mulheres.



**Figura 8** A ordem dos clíticos em construções com uma só forma verbal na fala do PM segundo Vieira (2002) – *Corpus PPOM*

Fonte: Vieira (2002: 85)

De modo geral, a distribuição das ocorrências também remete à esperada semelhança entre as variedades europeia e moçambicana. Entretanto, ao detalhar o comportamento dos dados quantitativa e qualitativamente, Vieira (2002) detecta particularidades do PM. Segundo a autora, a tendência do PM à variante pós-verbal é expressiva, ocorrendo, inclusive, em contextos em que se espera a anteposição do pronome ao verbo. Em outras palavras, a variedade moçambicana tende a realizar a ênclise de forma generalizada.

Esse comportamento geral foi atestado por pesquisadores (GONÇALVES; MORENO; TUZINE; DINIZ; MENDONÇA, 1998) que realizaram um primeiro diagnóstico das principais especificidades do Português Oral de Maputo, incluindo os produzidos no âmbito do *Corpus PPOM*. Considerando as estruturas gramaticais do Português que demonstram as áreas de dificuldades de aprendizes moçambicanos, os referidos autores declararam:

No que diz respeito aos padrões da ordem dos pronomes átonos, registam-se com maior frequência casos de adopção do padrão enclítico (*V-Pron*) (ver também Nhampule (1996) e Gonçalves (1997)), mesmo quando se trata de contextos que, de acordo com a norma europeia, exigem a próclise. (p. 62)

A fim de demonstrar os condicionamentos da próclise na variedade moçambicana, Vieira (2002) detalha o comportamento das variáveis que se mostraram

sistematicamente relevantes<sup>12</sup> ao condicionamento da ordem dos clíticos no *Corpus PPOM*, quais sejam: a *presença de operador de próclise na oração* e a *distância entre o operador de próclise e o grupo clítico-verbo*. Com o controle detalhado dos fatores da primeira variável, a análise empreendida pela autora forneceu os resultados detalhados na Tabela 5.

**Tabela 5** Aplicação da próclise de acordo com a variável *Presença de operador de próclise* em construções com uma só forma verbal no PM segundo Vieira (2002) – *Corpus PPOM*

| OPERADOR DE PRÓCLISE                 | VALOR ABSOLUTO | PERCENTUAL | PESO RELATIVO |
|--------------------------------------|----------------|------------|---------------|
| Nenhum                               | 8/285          | 3%         | .06           |
| SN sujeito nominal                   | 10/57          | 18%        | .38           |
| SN sujeito pronomes                  | 7/56           | 13%        | .30           |
| Conjunções coordenativas             | 12/125         | 10%        | .17           |
| SAdv – <i>aqui, ali, cá, lá, já</i>  | 16/28          | 57%        | .61           |
| SAdv – <i>sempre, talvez...</i>      | 9/31           | 29%        | .36           |
| SAdv – <i>-mente</i> e loc.adv.      | 1/15           | 7%         | .22           |
| Elemento “denotativo”                | 6/20           | 30%        | .34           |
| Preposição <i>a, em</i>              | 3/20           | 15%        | .36           |
| Preposição <i>para, de, sem</i>      | 113/138        | 82%        | .66           |
| Partícula de negação                 | 177/179        | 99%        | .99           |
| Conj. subordinativa e int. <i>se</i> | 56/79          | 71%        | .60           |
| Conjunção integrante <i>que</i>      | 25/52          | 48%        | .29           |
| Elemento <i>que</i> – outros         | 182/229        | 79%        | .73           |
| Pron./adv. relativo + pal. QU-       | 26/32          | 81%        | .74           |

Fonte: Vieira (2002: 191)

De acordo com os resultados, constituem elementos favorecedores da variação pré-verbal: a partícula de negação, os pronomes/advérbios relativos e palavra

<sup>12</sup> Vieira (2002) demonstra que atuaram também, em segundo plano, o *tipo de oração* e o *tipo de clítico*, variáveis que não puderam ser apresentadas considerando os propósitos e os limites do presente artigo.

QU-, o elemento *que* (excetuando-se a conjunção integrante), as preposições (exceto *a* e *em*), o SAdv do tipo *aqui* e a conjunção subordinativa/integrante *se*. Não favorecem a próclise os seguintes fatores: ausência de operador, conjunção coordenativa, SAdv do tipo *sempre*, SAdv em *-mentel* locução adverbial, SN sujeito nominal, elemento chamado denotativo (operador de foco), preposições *a* e *em* e conjunção integrante *que*. Levando em consideração todos os contextos possíveis reunidos ao máximo por semelhança, Vieira (2002) apresenta os resultados reunidos em quatro fatores, conforme se expõe na Tabela 6.

**Tabela 6** Aplicação da próclise de acordo com a *Presença de operador de próclise na oração* (variantes amalgamadas) em construções com uma só forma verbal no PM segundo Vieira (2002)

| OPERADOR DE PRÓCLISE                | VALOR ABSOLUTO | PERCENTUAL | PESO RELATIVO |
|-------------------------------------|----------------|------------|---------------|
| Nenhum                              | 8/285          | 3%         | .06           |
| SN sujeito                          | 17/113         | 15%        | .30           |
| Conj. coordenativa/ Exp. adverbiais | 38/199         | 19%        | .27           |
| Operadores diversos                 | 588/749        | 79%        | .81           |

Fonte: Vieira (2002: 192)

Não constituem, de um lado, contextos favorecedores da próclise a ausência de um operador (.06), o SN sujeito (.30), a conjunção coordenativa e as expressões adverbiais (.27); de outro lado, figuram os demais operadores, que favorecem a variante pré-verbal (.81).

O efeito da distância entre o operador de próclise e o grupo clítico-verbo também se confirmou, como se pode observar nos resultados expostos na Tabela 7.

**Tabela 7** Aplicação da próclise de acordo com a *Distância entre o elemento proclisador e o verbo na oração* em construções com uma só forma verbal no PM segundo Vieira (2002) – *Corpus PPOM*

| DISTÂNCIA ENTRE O OPERADOR E CL-V | VALOR ABSOLUTO | PERCENTUAL | PESO RELATIVO |
|-----------------------------------|----------------|------------|---------------|
| Zero sílaba                       | 542/726        | 75%        | .56           |
| Uma a cinco sílabas               | 66/156         | 42%        | .33           |
| Seis a dez sílabas                | 8/34           | 24%        | .13           |
| Onze sílabas em diante            | 9/18           | 50%        | .43           |

Fonte: Vieira (2002: 193)

A codificação dessa variável evidencia que, no PM, a distância atenua, até certo ponto, o efeito da operação de próclise. Essa afirmativa baseia-se na redução gradativa dos índices de próclise do contexto *zero sílaba* até *seis a dez sílabas* (.56 > .33 > .13). O último fator – *onze sílabas em diante* – apresenta índice de próclise (.43) semelhante ao da média geral (.49) do *corpus*. Parece que, estando o operador muito distante do clítico, o efeito da chamada atração deixa de existir, de modo que os dados se comportam conforme a tendência geral da variedade moçambicana.

Embora se tenha atestado a relevância dos grupos de fatores *presença de operador de próclise na oração* e *distância entre o operador e o grupo clítico-verbo*, o condicionamento dessas variáveis no PM é menos sistemático do que o verificado no PE: o comportamento dos fatores evidencia a já referida instabilidade da chamada operação de próclise. Tanto a seleção desses grupos quanto o fato de com eles ora interagirem outras variáveis de natureza morfossintática – como *tipo de oração*, *tempo e modo verbais* e *tipo de clítico* – revelam essa assistemática. A esse respeito, a observação de exemplos reafirma o verificado.

Ainda que não sejam de alta produtividade, chamam a atenção ocorrências de clíticos nos seguintes contextos:

\* em início de oração

Ex. 32. ((pergunta do doc.)) me é difícil porque o meu pai ainda não arranjou nenhuma madrasta de lá ate cá (PMA2-PPOM, PC1URA)

\* com SN sujeito nominal:

Ex. 33: os rapazes e as raparigas se preparam se conhecem se conhecem para que futuramente consigam viver em conjunto (PMA3-PPOM, inf. PC9GRI)

Ex. 34: quando eu nasci comecei a: na machamba a minha avó me obrigava muito para andar na machamba (PMC1-PPOM, inf. MX16LIN)

\* com SN sujeito pronominal:

Ex. 35: dezassete horas tinha que estar em casa... trancada... não podia brincar com ninguém... eu me sentia muito mal porque eu gostava de brincar (PMA2-PPOM, inf. PC1URA)

Ex. 36: partindo mesmo do namoro... ainda novos e: enquanto existe esse abuso.. essa falta de respeito em casa... eles se precipitam... “ah vamos casar” porque tem dinheiro tem condições “vamos casar” casam-se (PM-A1-PPOM, inf. MF13SUR)

\* após conjunções coordenativas e/ou advérbios em –mente:

Ex. 37: tem praticamente essa janela fechada e praticamente se torna ignorante – e a ignorância – é um mal extremamente grande (PM-A3-PPOM, inf. PC9GRI)

Ex. 38: mulheres mas se diz homens engenheiros e mulheres (PM-A3-PPOM, inf. PC19CUN)

Como se pode observar, mesmo na ausência de elementos proclisadores, foi registrada a variante proclítica, comportamento não aferido nas amostras do PE que foram estudadas até o momento.

De outro lado, diversas ocorrências de ênclise ocorreram em contextos com a presença de operadores de próclise. Segundo Vieira (2002), “na variedade oral moçambicana, o desrespeito à tradicional “atração” alcança um número considerável de dados, especialmente no caso da conjunção integrante <que>, que chega a registrar mais ênclise do que próclise (48%)” (VIEIRA, 2002, p. 111). A autora oferece diversos exemplos para comprovar essa tendência.

Ex. 39: a falta de respeito posso não posso dizer que deve-se ao professor porque muitas das vezes até os alunos faltam respeito ao professor (PM-A3-inf. AM23VRA) [conjunção integrante que]

Ex. 40: os médicos diziam que eu sentia-me fraca – cansaço (PMA1-PPOM, inf. MX3MAR) [conjunção integrante que]

Ex. 41: agora eu não sei a: a razão principal se de facto deve-se à falta de namoro (PMA2-PPOM, inf. PC6GRA) [conjunção integrante se]

Ex. 42: sempre a do noivo é de dizer <<não tenho dinheiro porque todo o dinheiro gastei no lobolo>> e que só depois de muito tempo ou chegasse a não... a não realizar-se o casamento – porque ele sempre vai-se desculpar porque não tem dinheiro (PMA2-PPOM, inf. CH12ODJ) [partícula de negação]

Ex. 43: não posso contar assim muito... sei lá... porque sempre quando desloco-me a Nampula o meu pai tem ido lá em missão de serviço (PMA2-PPOM, inf. CH12ODJ) [conjunção subordinativa]

Ex. 44: a igreja é muito importante – muito importante eu acho porque: nasci e cresci na igreja até agora – embora diz-se – né? haver certas igrejas que existe só para explorar ou apoderar-se do dinheiro de alguns (PMA1-PPOM, inf. MX2SIM) [conjunção subordinativa]

Ex. 45: acho que não tem possibilidade para tal: mas se alguém abre-se – amostra a simpatia dele – que é a educação dos próprios pais – deve-te cumprimentar (PMA1-PPOM, inf. MX3MAR) [conjunção subordinativa]

Ex. 46: depende da sorte que elas tiverem né? do marido ou namorado que apareceu-lhes em frente (PM-A2-PPOM, inf. CH12ODJ) [pronome relativo que]

Ex. 47: isso não sei dizer porque os meus pais é que disseram-me (PMB1-PPOM, inf. MF1ANA) [estrutura clivada]

Ex. 48: a minha infância direi que uma parte não foi feliz a outra foi – não foi feliz porque eu era muito mimado com os pais – não deixavam-me sair tanto de casa – como/ quem batia-me fora sempre eles iam lá defendiam-me (PMA3-PPOM, inf. PC2CEL) [pronome relativo quem]

Ex. 49: porque tínhamos que dormir no mato tinha mesmo sítio próprio onde a gente tinha que dormir onde chamava-se de placa eram sítios que – basta anoitecer tínhamos todos que correrem (PMB1-PPOM, inf. MF1ANA) [advérbio relativo onde]

Com base nos índices quantitativos e na análise qualitativa das ocorrências, a autora chega a concluir sua apreciação da seguinte forma:

Ao que parece, esse número expressivo de casos em que o possível operador de próclise não “opera” revela um traço caracterizador da variedade moçambicana e não uma espécie de “deslize” em relação ao que seria “normal” ocorrer. (VIEIRA, 2002, p. 111)

A esse respeito, Gonçalves (2010) generaliza e confirma o que se registrou anteriormente em Gonçalves; Stroud (1998), mencionando ser a colocação pronominal fenômeno sintático que distingue PM e PE.

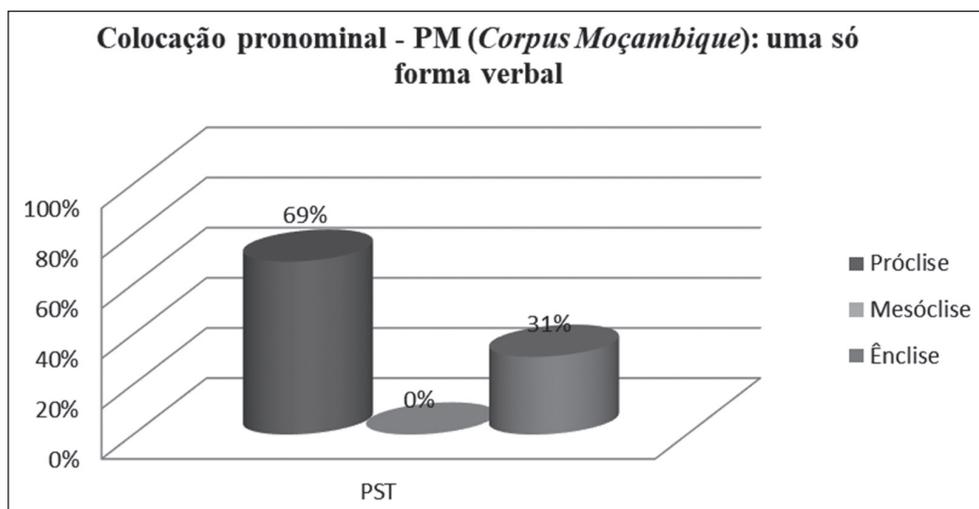
Quanto aos fenômenos mais estritamente sintáticos que distinguem o PM do PE, e no que diz respeito aos padrões de ordem dos pronomes pessoais átonos, salienta-se em primeiro lugar a tendência para adotar a ênclise em contextos em que estão presentes <<atratores>> de próclise, sobressaindo, pela sua frequência, os casos em que está presente um complementador. (GONÇALVES, 2010, p. 52)

A fim de observar se as tendências ora descritas no *Corpus PPOM* se verificam em outros materiais, apresentam-se, a seguir, resultados preliminares da análise das entrevistas do *Corpus Moçambique*<sup>13</sup>. Leva-se, aqui, em consideração

<sup>13</sup> Está em curso a análise dos 18 informantes da amostra de entrevistas sociolinguísticas no âmbito do Projeto *Estudo comparado dos padrões de concordância em variedades africanas, brasileiras e europeias do Português*. Os resultados dessa análise permitirão maior detalhamento não só das construções com uma só forma verbal, mas também com os complexos verbais.

a fala de apenas seis entrevistas sociolinguísticas produzidas em 2016, referentes a informantes jovens (de 18 a 35 anos de idade), distribuídos por sexo e escolaridade, sendo dois de cada nível – fundamental, médio e superior.

No referido conjunto de entrevistas, foram coletados 153 clíticos em construções com uma única forma verbal, com a distribuição pelas variantes controladas conforme representada na Figura 9.



**Figura 9** A ordem dos clíticos em construções com uma só forma verbal na fala do PM – *Corpus Moçambique*

Comparando esses resultados gerais aos obtidos por Vieira (2002), verifica-se, em termos absolutos, índice menor do uso dos pronomes oblíquos na posição pós-verbal. Os índices revelaram alta produtividade da próclise na amostra observada, que chega a 69,3% dos dados coletados na fala dos seis informantes.

Da análise variacionista, foram consideradas relevantes para o condicionamento da próclise as variáveis *Escolaridade*, *Sexo*, *Língua 1 do informante* e *Presença de elemento proclisador*.

Em termos linguísticos, embora os presentes resultados sejam preliminares, eles confirmam a particularidade nos padrões de colocação pronominal da variedade moçambicana. Em alguns contextos, ao que tudo indica, o PM segue as tendências conhecidas na variedade europeia (como o efeito proclisador ou não proclisador, de modo geral); em outros, chama atenção o registro de ocorrências que não costumam ser observadas no PE, mas são comumente encontradas na variedade brasileira. A Tabela 8 apresenta os resultados obtidos para a variante pré-verbal em cada contexto controlado.

**Tabela 8** Aplicação da próclise em construções com uma só forma verbal de acordo com a variável *Efeito proclisador* no PM – *Corpus Moçambique*

| EFEITO PROCLISADOR                     | V. ABSOLUTO | PERCENTUAL | PESO RELATIVO |
|--|-------------|------------|---------------|
| Início absoluto                        | 3/20        | 15%        | .06           |
| Sujeito SN                             | 8/17        | 47.1%      | .10           |
| Sujeito – Pronome                      | 17/18       | 94.4%      | .73           |
| Advérbio                               | 6/11        | 54.5%      | .29           |
| Locução adverbial e advérbio em -mente | 0/5         | 0%         | –             |
| Conjunção coordenativa                 | 7/10        | 70.0%      | .29           |
| Preposição                             | 14/15       | 93.3%      | .78           |
| Elemento subordinativo                 | 30/32       | 93.8%      | .83           |

Em primeiro lugar, chama atenção a presença de próclise em quase todos os contextos controlados (exceto após *locução adverbial/advérbio em -mente*, contexto com apenas cinco dados, todos em ênclise), inclusive em início de oração, embora este tenha se mostrado como contexto altamente favorecedor da variante pós-verbal, como se atesta nos exemplos 50 e 51.

Ex. 50: é muito comum música angolana ouvir música angolana cá... mas também já vai crescendo o número de música moçambicana... ou-ve-se pouco é verdade...também devido à falta de... de qualidade... mas já: há coisas de qualidade... na praça (PMA3H-Corpus Moçambique)

Ex. 51: imagine por exemplo num casamento... são dois normalmente na cultura ocidental... eh o eh eh eh... vai-se à igreja ou ou ou registo né... casam... (PMA3H-Corpus Moçambique)

Quanto aos demais contextos, é possível delimitar o conjunto dos elementos que favorecem a próclise e os que não a favorecem. Pertencem ao primeiro grupo os seguintes fatores: *elemento subordinativo*, *sujeito na forma pronominal* e *preposição*. Pertencem ao grupo de elementos desfavorecedores da próclise os seguintes fatores: *início absoluto*, *sujeito na forma nominal*, *advérbio*, *locução adverbial/advérbio em -mente* e *conjunção coordenativa*.

Há que se destacar o número expressivo de dados de próclise em contextos diferentes dos verificados para a variedade europeia. A esse respeito, observam-se, particularmente, os dados em que os clíticos são antecidos de pronomes na função de sujeito, os quais registraram 94.4% de próclise (.73), tendo sido produzidos por indivíduos de diferentes níveis de escolaridade.

Ex. 52: eu encontrei um senhor/ isso qual foi em dois mil e onze senhor falou para mim disse assim “peço ajuda” era no domingo eu “ajuda sim” “trago aqui pedras preciosas” “eu preciso de chegar no câmbio” ele **me amostrou** (PMA1H-Corpus Moçambique)

Ex. 53: e: educação saúde quer dizer com o nível de vida que levamos que não é bom hhh nós **nos forçamos** muito (PMA3M-Corpus Moçambique)

Quanto aos demais contextos favorecedores da próclise (preposição e elementos subordinativos), os resultados apenas confirmam o verificado em amostras da variedade europeia da Língua Portuguesa. No que diz respeito à preposição, o PM realizou a próclise segundo uma taxa de 93.3% (.78); o elemento subordinativo também apresentou preferência pela variante pré-verbal, com taxa de 93.8% (.83), confirmando a tendência à variante pré-verbal, embora também não seja de forma categórica.

Em termos extralinguísticos, o estudo preliminar da fala dos informantes mais jovens do *Corpus* Moçambique sugere que a preferência pela próclise parece estar correlacionada à fala das mulheres, ao aprendizado do Português como segunda língua e ao aumento da escolaridade.

Primeiramente, ao que tudo indica, o controle do uso da variante pré-verbal respeitando os contextos em que ela seria preferida na variedade europeia decorreria do maior contato e aprendizagem do Português. O grau do conhecimento formal das regras influenciaria diretamente na fala.

**Tabela 9** Aplicação da próclise em construções com uma só forma verbal de acordo com a variável *Escolaridade* no PM – *Corpus Moçambique*

| ESCOLARIDADE       | VALOR ABSOLUTO | PERCENTUAL | PESO RELATIVO |
|--------------------|----------------|------------|---------------|
| Ensino fundamental | 35/49          | 71.4%      | .51           |
| Ensino médio       | 3/59           | 72.9%      | .69           |
| Ensino superior    | 28/45          | 62.2%      | .24           |

De acordo com os resultados obtidos, o nível de escolaridade elevado faz com que o falante produza menos a próclise. Em termos relativos, aumenta a tendência à próclise do fundamental (.51) para o médio (.69), enquanto o curso superior funciona como um dos elementos que desfavoreceria a variante (.24). Em termos qualitativos, espera-se que estruturas com próclises em contextos sem elementos proclisadores seriam mais prováveis na fala de indivíduos menos escolarizados, como o que produziu o exemplo 54.

Ex. 54: eu acho que eu poderia sair eu ia ficar na minha casa eu ia criar meus filhos como minha mãe fez... minha mãe **nos criou** independentemente dela sem condições (PMA1M-Corpus Moçambique)

De outro lado, exemplos como o 55 reproduzem a preferência pela ênclise em contextos de clíticos antecidos de SN sujeito na fala dos informantes com curso superior. Supõe-se, assim, que o conhecimento formal das regras pode influenciar a escolha da colocação pronominal.

Ex. 55: as meninas **prostituem-se** ou se envolvem com pessoas mais velhas e os rapazes **metem-se** à criminalidade juntam-se aos grandes mercados informais (PMA3M-Corpus Moçambique)

A variável *Sexo* também demonstrou influência sobre o fenômeno no *Corpus Moçambique*, conforme os índices expostos na Tabela 10.

**Tabela 10** Aplicação da próclise em construções com uma só forma verbal de acordo com a variável *Sexo do informante* no PM – *Corpus Moçambique*

| SEXO DO INFORMANTE | VALOR ABSOLUTO | PERCENTUAL | PESO RELATIVO |
|--------------------|----------------|------------|---------------|
| Homem              | 40/65          | 61.5%      | .30           |
| Mulher             | 66/88          | 75%        | .64           |

Os resultados ora apresentados apontam o favorecimento da próclise pelas mulheres (.64), em comparação aos homens (.30). Conforme já se advertiu quanto aos resultados obtidos para o PST, a interpretação do comportamento dessa variável não pode ser seguramente feita sem aprofundamento quanto ao papel da mulher na sociedade em questão. Em termos sociais, a mulher da capital Maputo, ao que parece, é bastante ativa no mercado de trabalho, assumindo, muitas vezes, a responsabilidade financeira da casa; ocorre que só um estudo aprofundado do perfil feminino permitirá investigar o comportamento da variável. Ademais, é preciso aprofundar a análise observando os contextos morfossintáticos em questão do ponto de vista estrutural: de um lado, a preferência da próclise pelas mulheres pode estar relacionada ao maior grau de atenção ao efeito das partículas proclisadoras; de outro, esse resultado pode se relacionar ao não domínio das regras estruturais e à produção da variante pré-verbal independentemente dos chamados fatores de próclise.

Além da escolaridade e do sexo dos informantes, a variável referente ao *Estaduto da Língua Portuguesa* para o falante – se ele a declara como sua língua materna ou não – também se mostrou relevante ao condicionamento da ordem dos clíticos pronominais. Vieira (2002) atribuiu, por hipótese, certa tendência à

generalização da ênclise na variedade moçambicana ao aprendizado do Português como segunda língua.

**Tabela 11** Aplicação da próclise em construções com uma só forma verbal de acordo com a variável *Estatuto da Língua Portuguesa* no PM – *Corpus Moçambique*

| ESTATUTO DA LÍNGUA PORTUGUESA | VALOR ABSOLUTO | PERCENTUAL | PESO RELATIVO |
|-------------------------------|----------------|------------|---------------|
| Primeira língua               | 84/106         | 79.2%      | .59           |
| Segunda língua                | 22/47          | 46.8%      | .29           |

Embora não seja possível propor generalizações, dada a pequena quantidade de dados ora em análise, chama a atenção a seleção dessa variável como relevante ao condicionamento do fenômeno. Segundo os resultados obtidos, os falantes de Português como língua materna realizaram próclise em 79.2% dos dados (.59), enquanto os informantes que se caracterizaram como falantes de Português como segunda língua realizaram a variante em 46.8% dos dados (.29), o que sugere desfavorecimento da próclise e possível uso mais geral de ênclise, conforme se observou no comportamento da variável *Presença de elemento proclisador*.

#### 4. O ESTATUTO DA ORDEM DOS CLÍTICOS NAS VARIEDADES DA LÍNGUA PORTUGUESA

Com base em diversos trabalhos a respeito do tema da colocação pronominal e nos resultados apresentados no presente artigo, é possível sistematizar os resultados e refletir sobre a possível interpretação quanto ao perfil gramatical de cada variedade do Português sob análise.

No que se refere às construções com uma só forma verbal, o PE apresenta a variante enclítica nos contextos *início absoluto de oração*, após *sujeito*, *conjunções coordenativas*, *expressões adverbiais*, dentre outros elementos não proclisadores, mas a variante pré-verbal é a mais produtiva – embora não categórica – nos casos em que há elementos proclisadores. Nos dados do PB, do PST e do PM, entretanto, verificam-se ao menos dois padrões diferentes de colocação dos clíticos pronominais: enquanto no PB a regra é semicategórica, verifica-se, no PST e no PM, uma regra variável com graus de aproximação ao modelo europeu distintos.

No PB, não há efetiva variação, uma vez que a colocação foi predominantemente proclítica, independentemente do ambiente sintático em que o clítico se encontra. Os poucos dados de ênclise ocorreram com determinados pronomes átonos (sobretudo *o* e variantes, e *se* indeterminador) e em estruturas específicas.

A variedade são-tomense utiliza, em contextos não marcados, mais a variante enclítica do que a proclítica. Destaca-se que a colocação do clítico no início absoluto de oração foi categoricamente pós-verbal, igual à registrada no PE. Os resultados acompanham, em geral, a tendência europeia de colocação, mas não seguindo com o mesmo rigor as regras que imporiam a ordem pré-verbal. Assim, não há uma nítida dicotomia entre elementos proclisadores *versus* elementos não proclisadores, como ocorre no PE (com certa variação), mas certa oscilação na colocação em mesmos contextos sintáticos. Além disso, a colocação proclítica ocorreu, por vezes, com elementos antecedentes não considerados proclisadores (como tipicamente ocorre no PB), como, por exemplo, com sujeitos e conjunções coordenativas.

No PM, de forma semelhante ao registrado no PST, também se constata uma regra variável, com preferência pela ênclise em contextos não marcados e a atuação de certas partículas como proclisadoras. Nesse quadro geral, entretanto, entende-se que a instabilidade no comportamento da regra variável na variedade moçambicana é ainda maior do que a registrada na são-tomense. Ao mesmo tempo em que o efeito proclisador é verificado nos dados, é fato que se registram ocorrências não observadas – como a próclise em início de oração – ou não produtivas – como após sujeitos pronominais, por exemplo – no PE.

No que diz respeito aos complexos verbais, a análise dos dados do PE revela que há variação nos complexos formados pelas três formas de verbo principal, não tendo ocorrido a efetiva próclise a v2. Em complexos com infinitivo, ocorrem as posições proclítica a v1, enclítica a v1 e a v2, enquanto, no gerúndio e no particípio, a variação foi registrada somente entre as duas primeiras, a próclise ou a ênclise a v1.

Em relação às demais variedades – neste caso, foram considerados apenas PB e PST –, há tipos de regra específicos para cada uma delas.

Em dados brasileiros de complexos verbais, assim como nas construções com uma só forma verbal, não se verificou efetiva regra variável, tendo em vista que a colocação entre as duas formas verbais – interpretada, aqui, como próclise a v2, em função do comportamento assumido nos contextos com elementos intervinientes – constitui a opção preferencial. Os poucos dados com a colocação proclítica a v1 e enclítica a v2, sobretudo nos complexos com infinitivo, ocorrem apenas com determinados pronomes e estruturas.

No PST, ocorrem as mesmas posições realizadas no PE com os verbos no infinitivo (próclise ou ênclise a v1 e ênclise a v2), gerúndio e particípio (proclítico ou enclítico a v1), mostrando variação com os três tipos de verbos principais. De todo modo, chama atenção a ocorrência de clíticos adjacentes a v2 não somente após *ter que/de*, como ocorre no PE, mas também após a preposição *a* e o vocá-

bulos *até* entre as duas formas verbais, estruturas cuja explicação merece ser aprofundada em estudos futuros que permitam observar maior número de dados com complexos verbais e em maior diversidade de estruturas.

A partir dos resultados ora sintetizados, seja nas construções com uma só forma verbal, seja nos complexos verbais, há que se refletir sobre a formação de cada variedade do Português:

(i) O Português do Brasil

Na variedade brasileira, cujos padrões gramaticais já estão sedimentados ou fixados, trabalhos de natureza diacrônica (cf. PAGOTTO, 1992; MARTINS, 1994; MARTINS, 2009; e NUNES, 2009; 2014) demonstram que, durante todo o processo de edificação da língua, foram observados diversos momentos de variação. Assim, no século XVI, época em que os portugueses chegaram ao Brasil, constituiria a opção preferencial a realização da variante proclítica (em todos os contextos, exceto em início absoluto de oração); no século XVII, teria começado uma mudança nos padrões de colocação pronominal, de modo que PE e PB tomaram rumos diferenciados. Resultou dessa história que, enquanto o PE passou a restringir a variante proclítica aos contextos com proclisador, o PB continuou utilizando a variante proclítica e até generalizou seu uso, empregando-a inclusive em início absoluto de oração.

Muito debate já se empreendeu a respeito das motivações para essa alteração gramatical, dentre as quais avulta a hipótese relacionada aos padrões rítmicos brasileiros, em oposição aos portugueses, de modo que o clítico pronominal, fonologicamente, seria uma partícula que se apoiaria, em termos de ligação fonológica, a enunciados tônicos à direita, enquanto no PE o padrão ocorreria em direção à esquerda (cf. VIEIRA, 2002; CORRÊA, 2012; VIEIRA; CORRÊA, 2017). Essa suposta ligação do PB para a direita fica bem nítida em dois aspectos: (i) no fato de haver próclise em início absoluto na variedade brasileira e após elementos não proclisadores; e (ii) no caso dos complexos verbais, no fato de haver a próclise a v2.

De acordo com diversos estudos sobre o tema<sup>14</sup>, estaria na base do encaixamento dessa mudança o fato de o PB ter se diferenciado do PE em função da redução das vogais átonas europeias (século XVII) com consequentes alterações no padrão rítmico. Assim, as partículas átonas desencadeariam a cliticização à esquerda, fazendo com que a ênclise se tornasse a colocação mais geral e a variante pré-verbal se apoiasse igualmente à esquerda nos chamados

---

<sup>14</sup> Motivações extralinguísticas para essa mudança gramatical não estão efetivamente descartadas, mas carecem de resultados científicos até o momento.

elementos proclisadores<sup>15</sup>. No PB, no entanto, variedade em que essa redução das vogais átonas não ocorreu, a cliticização fonológica do pronome átono fortaleceria sua ligação à direita. Dessa forma, a próclise pôde ser generalizada, passando a ocorrer inclusive em início absoluto de oração. O parâmetro de colocação pronominal brasileiro resultante de todo esse processo pode ser verificado claramente nos resultados sintetizados no presente trabalho, que mostrou que não há efetivamente variação na modalidade oral, uma vez que a regra se revelou semicategórica com preferência pela próclise em todos os contextos.

#### (ii) O Português de São Tomé

No que se refere à variedade são-tomense, os padrões locais da Língua Portuguesa – embora utilizada como L1 pela maioria das pessoas – ainda não parecem estar totalmente consolidados. É certo que o intenso contato interlinguístico, em especial a convivência com um dos crioulos de base portuguesa, o Forro, também utilizado pela maioria da população, colabora para essa instabilidade na fixação dos padrões de colocação pronominal.

Em termos de tendências gerais, é nítida a preferência pelos padrões europeus de colocação, sendo a ênclise categórica, por exemplo, em início absoluto de oração. Embora a colocação seja sensível à ação de elementos no contexto antecedente ao clítico, não há uma efetiva polaridade entre elementos proclisadores *versus* não proclisadores; o que há é uma tendência à variante pré-verbal mediante a atuação de proclisadores, mas também a realização da ênclise nos mesmos contextos sintáticos, situação que acontece em dados do PE em proporções comparativamente muito reduzidas. Associados a essa instabilidade da atuação dos elementos proclisadores, registram-se alguns dados compatíveis com os usos brasileiros, como, por exemplo, o uso de próclise após sujeito e conjunções coordenativas.

Com os complexos verbais, essa oscilação também ocorre, porém de forma bem menos evidente. Embora, de modo geral, as tendências também sejam semelhantes às verificadas no PE, chama atenção, nos contextos com elementos proclisadores, a variação entre ocorrências de próclise a v1, ênclise a v1 e a v2 – enquanto no PE a alternância ocorreu prototipicamente apenas entre próclise a v1 e ênclise a v2. Ademais, algumas poucas ocorrências de clítico após a preposição *a*, em estruturas aspectuais, e uma única após *até* – o que não se registra no PE –

<sup>15</sup> Segundo Martins (1994), essa alteração paramétrica no PE estaria relacionada à chamada categoria sigma. Nesse nível sintático, haveria inicialmente a regulação da modalidade frasal, negativa ou afirmativa, o que se correlacionaria a alterações relacionadas a outras opções gramaticais, segundo o aporte da teoria gerativa, e, em última análise, afetaria a ordem dos clíticos.

deixam dúvida quanto à possibilidade de o clítico do PST estar associado em próclise a v2.

(iii) O Português de Moçambique

Considerando um *continuum* de maior ou menor proximidade dos dados com o PB ou com o PE, a variedade moçambicana permite postular maior afastamento da norma que teoricamente lhe serve de referência, a europeia.

Em termos linguísticos, embora os presentes resultados sejam preliminares, eles sugerem uma particularidade ainda mais clara no quadro estrutural da Língua Portuguesa em formação na variedade moçambicana. Aparentemente, a variedade em questão estaria estabelecendo suas feições ancoradas na instabilidade de colocação, visto que, em alguns contextos, segue as tendências conhecidas na variedade europeia (como o efeito proclisador ou não proclisador, de modo geral) e em outros chega a registrar ocorrências não observadas na variedade europeia (nem na são-tomense), mas típicas da variedade brasileira, como a presença de próclise em início de oração ou a preferência absoluta por essa variante após sujeitos pronominais, por exemplo. A força dos condicionamentos extralinguísticos, como os relacionados a escolaridade, sexo, estatuto da língua materna, também advoga em favor da especificidade do PM.

A dimensão da instabilidade moçambicana foi registrada no estudo pioneiro de SEMEDO (1997), quando atesta, na sintaxe moçambicana, “entre outros aspectos, uma grande liberdade no uso dos pronomes pessoais objecto átonos (clíticos).” (p. 8). O autor propõe que, embora o Português utilizado em Moçambique siga, oficialmente, a norma europeia, ocorrem particularidades evidentes:

no que concerne ao domínio de colocação dos pronomes pessoais clíticos, os inquiridos seguem a norma do português padrão em 50% dos casos, enquanto o grupo de controle [falantes lisboetas] o faz em 92% (...) (SEMEDO, 1997, p. 32).

Embora a investigação dos dados de entrevistas recentes, como as do *Corpus Moçambique*, ainda esteja em estágio inicial, o que se verificou nas análises apresentadas pode ser um possível reflexo da particular situação de multilinguismo e do estatuto do Português como L2 em Moçambique, apontando a configuração de padrões próprios dessa variedade. Assim, os resultados ora apresentados indicam a necessidade de estudos aprofundados dos fenômenos variáveis em uso por moçambicanos, de modo a conseguir identificar raízes das possíveis mudanças e sinais de que essa variedade do Português já esteja prestes a assumir um perfil próprio, assim como aconteceu com a variedade brasileira.

Os resultados do estudo vão ao encontro da hipótese de que o estatuto de língua aprendida (L2) do PM para muitos falantes (em uma sociedade cuja histó-

ria de opressão se ressentia da defesa de seus próprios elementos culturais) possa estar relacionado ao caráter hesitante da expressão linguística. Essa espécie de ‘conflito’ no uso da língua foi apreendida, com mestria, pelo escritor moçambicano, em suas sábias palavras:

O português moçambicano – ou ainda, nesta altura, o português em Moçambique – é ele próprio um lugar de conflitos e ambiguidades. A adesão moçambicana à lusofonia está carregada de reservas, aparentes recusas, desconfiadas aderências. O que eu gostaria de mostrar aqui é esse caminho em zig-zag que não resulta de capricho dos dirigentes, mas de ambivalências da História. (MIA COUTO, 2002)<sup>16</sup>

## 5. REFLEXÕES FINAIS: AS PARTICULARIDADES DAS VARIEDADES AFRICANAS DO PORTUGUÊS

Com base na descrição dos dados do Português de São Tomé e do Português de Moçambique, cabem algumas reflexões finais a respeito das particularidades da colocação pronominal verificadas nessas variedades.

De um lado, o forte alinhamento das variedades africanas com os padrões de colocação europeus pode ser claramente relacionado à influência direta do processo de colonização tanto em São Tomé e Príncipe quanto em Moçambique. Desse modo, é mais do que natural que a norma que venha a servir de referência, em ambos os casos, seja a europeia, sobretudo se for considerada a situação de aprendizagem formal de Português como L2 nos bancos escolares. A esse respeito, é preciso, sem dúvida, investir em pesquisas que investiguem o perfil do professor, do material didático escolar e dos métodos de ensino de Português como L2, para que seja possível aquilatar melhor a influência do PE na aprendizagem formal da língua.

De outro lado, as particularidades verificadas no PST e no PM – dentre as quais se destacam, no âmbito do presente artigo, as relacionadas à ordem dos clíticos pronominais – necessitam ser explicadas à luz de hipóteses de ordem linguística e/ou extralinguística.

As motivações internas à própria língua para justificar os padrões de colocação pronominal diferentes entre PB e PE foram propostas mediante a soma de resultados de vasta pesquisa sincrônica e diacrônica, o que não seria possível

<sup>16</sup> As palavras ora citadas constituem trecho de sua palestra, intitulada “Luso-afonias – a lusofonia entre viagens e crimes”, proferida no IV Seminário das lusografias, realizado em Maputo, de 18 a 22 de fevereiro de 2002, a ser publicada nas *Actas do IV Seminário das Lusografias*.

propor nos limites do presente artigo. Ainda assim, essas explicações ficaram restritas a motivações de ordem sintática e de natureza rítmica.

Considerando, então, motivações de natureza externa, ao menos duas hipóteses explicativas podem ser apresentadas: (i) as particularidades verificadas adviriam da influência direta de um substrato; e (ii) as particularidades resultariam da situação de intenso contato linguístico típico das sociedades em questão, o que acabaria por gerar *inputs* distintos para aquisição do Português.

No caso das variedades africanas, a falta de conhecimento aprofundado do Forro (língua crioula mais falada em São Tomé) e do Changana (língua banto mais falada em Maputo) não permite, nos limites deste artigo, maior debate acerca da primeira dessas hipóteses.

Vieira (2016), entretanto, de posse da descrição de características do Forro, de acordo com Ferraz (1979) e Hagemeyer (2007), afirma que o crioulo utilizaria os pronomes objetos, geralmente, na posição pós-verbal. Assim sendo, a influência do Forro não explicaria as diversas ocorrências da variante pré-verbal no PST. Desse modo, a motivação da instabilidade na colocação dos clíticos verificada no PST e, de forma ainda mais intensa, no PM, embora pareça estar relacionada à situação de intenso contato com diversas línguas, carece de investigação aprofundada e específica.

Sem dúvida, os resultados dos estudos reunidos no presente texto somam-se à interpretação feita por pesquisadores que têm defendido a tese de que o contato entre línguas distintas constitua forte motivação para as configurações das variedades não nativas de línguas coloniais (LUCCHESI; BAXTER; RIBEIRO, 2009, por exemplo). A respeito do comportamento específico do PST, declara, também, Gonçalves (2009):

(...) a variação entre próclise e ênclise nos mesmos contextos sintáticos é frequente no Português Oral de S. Tomé. Este facto é típico de uma variedade de língua que, embora actualmente comece a afirmar-se como língua materna, foi adquirida como língua segunda. (...) Aquilo que os dados dos clíticos nos mostram é que, muito provavelmente, o português de S. Tomé constitui uma interlíngua ainda não fossilizada pelo que a variação nas mesmas estruturas não nos permite caracterizar o sistema de colocação dos clíticos mas tão-somente descrevê-lo o mais pormenorizadamente possível procurando entrever aquilo que pode vir a constituir uma idiosincrasia do seu sistema linguístico. (GONÇALVES, 2009, p. 22)

Assim, no caso do PST, a oscilação presente nos mesmos ambientes sintáticos seria favorecida por dois fatores: (i) de um lado, as influências para a constituição do PST favoreceriam a opção pós-verbal, seja pelo contato com o Forro,

direta ou indiretamente, seja pelo modelo europeu de colocação, que apresenta a ênclise como ordem não marcada; e (ii) de outro lado, pelo fato de ainda não se ter uma variedade são-tomense consolidada, a aquisição do modelo com efeito proclisador ainda não estaria estabelecida plenamente, de modo que ocorreria ora próclise, ora ênclise nos mesmos contextos sintáticos.

Estendendo a reflexão aos dados do PM, que se afastariam, ainda mais, das regras europeias de colocação, propõe-se que as tendências verificadas quanto à posição dos clíticos pronominais estejam vinculadas à própria história de formação dessa variedade, ainda em curso. Ao que tudo indica, o PM adotou inicialmente a opção preferencial da língua do colonizador: a ênclise como ordem não marcada, bem como a maioria dos condicionamentos sintáticos do PE; entretanto, não se submeteu plenamente às mesmas regras, havendo, assim, a referida instabilidade na utilização dos clíticos pronominais.

Comparando PM e PST, a maior ou menor distância das regras de uso de cada variedade em relação ao modelo europeu também pode estar relacionada ao estatuto do Português em cada uma dessas variedades. Sendo o Português L1 para a maioria dos falantes são-tomenses e L2 para a maioria dos moçambicanos, justificam-se as maiores particularidades bem como maior oscilação no uso da ordem dos clíticos no PM do que no PST, estando o primeiro, ao que parece, mais propenso à implementação da variante pré-verbal em contextos diversos.

A respeito da influência do contato multilinguístico e da aprendizagem instável do Português, convém uma reflexão adicional: considerando o Princípio de Uniformitarismo, o que a realidade do PST e do PM poderia revelar sobre a formação do PB? Como no Português do Brasil, na época da formação de suas especificidades, havia também diversos parâmetros linguísticos em contato, é bastante razoável propor que as características, sobretudo as relacionadas à instabilidade da formação da nova língua, eram semelhantes – resguardadas as dimensões – às verificadas nas realidades são-tomense e moçambicana.

Ao que parece, a situação de multilinguismo no Brasil pode ter exercido fundamental influência quanto aos padrões de colocação pronominal que, ao longo do tempo, resultaram em forte parâmetro diferenciador entre PB e PE. A próclise, recebida na colonização, teria se fixado como tendência, seja por possíveis padrões rítmicos ou estruturais relacionados às línguas bantas aqui faladas – o que o presente trabalho não pode afirmar, visto que não desenvolveu esse tópico como objeto de estudo –, seja pela instabilidade natural do comportamento linguístico em situações de intenso contato. Em diferentes medidas e direções, motivações sócio-históricas semelhantes podem estar relacionadas ao fato de que a língua dos países colonizados se distancie e utilize parâmetros de colocação diferentes da língua do país colonizador. Para que essa interpretação seja seguramente tratada,

entende-se ser necessária uma investigação específica e detalhada, como o fez Ferraz (1979) para o Forro, das línguas faladas em Moçambique e em São Tomé, e das que existiam no Brasil e que conviveram durante muitos séculos com a Língua Portuguesa.

De todo modo, as descrições da ordem dos clíticos pronominais no PB, PE, PST e PM ora propostas permitem delinear as diferenças entre as quatro variedades, que já podem ser percebidas. Primeiramente, a oposição mais evidente é a do PB contra o PE, diferença quantitativa e qualitativamente comprovada. Nesse sentido, particularizam o primeiro os contextos de início absoluto de oração – com próclise – e de complexos verbais – com próclise a v2, permitindo estabelecer a adoção de um parâmetro absolutamente particular para a variedade brasileira. Em segundo plano, o PST e o PM – em diferentes medidas – particularizam-se quanto às demais variedades por, de um lado, se assemelharem ao PE, adotando preferencialmente as mesmas variantes nos contextos aqui destacados e aplicando tendências de uso consoante o efeito proclisador; de outro lado, entretanto, apresentam quantitativamente uma escala gradual desse efeito e não uma efetiva delimitação dos contextos de próclise e de ênclise. Ademais, qualitativamente, verificam-se dados de próclise nas duas variedades africanas em contextos que seriam de ênclise na variedade europeia.

Embora ainda sem condições de propor generalizações seguras, o presente artigo reafirma o necessário caminho para ampliar a reflexão sobre a formação do PB e das demais variedades do Português, que se afigura como de extrema importância. Sem dúvida, o estudo da colocação pronominal e de outros temas em diversas variedades da Língua Portuguesa, em especial nas que ainda estão, em diferentes medidas, em constituição, como o Português de São Tomé, de Moçambique, de Angola, entre outras, é fundamental para a compreensão do que ocorre não só com as variedades transplantadas, mas também, indiretamente, do que supostamente ocorreu com a formação do Português do Brasil. Desse modo, a proposta de um *continuum* afro-brasileiro possivelmente será sustentada, não como uma tese aprioristicamente definida, mas como uma consequência natural do conhecimento acerca das variedades do Português em relação a cada regra linguística, e em diversos níveis gramaticais, empreendimento de que resulta a publicação do presente livro.

## REFERÊNCIAS

- AVELAR, Juanito; GALVES, Charlotte. O papel das línguas africanas na emergência da gramática do português brasileiro. *Linguística*, 30 (2), p. 241-288, 2014.
- BRANDÃO, Silvia Figueiredo. Concordância nominal em duas variedades do Português: Convergências e divergências. *Revista Veredas*, 15 (1), p. 164-178, 2011.

CORRÊA, Cristina Márcia Monteiro de Lima. *Cliticização pronominal na região metropolitana do Rio de Janeiro: a interface Sintaxe-Fonologia*. 2012. Dissertação (Mestrado em Letras Vernáculas) – Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2012.

FERRAZ, Luis Ivens. *The Creole of São Tomé*. Johannesburg: Witwatersrand University Press, 1979.

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

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

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

GONÇALVES, Perpétua; MORENO, Albertina; TUZINE, António; DINIZ, Maria João; MENDONÇA, Marisa. *Estruturas gramaticais do Português: problemas e exercícios. Pronomes pessoais*. In: GONÇALVES, Perpétua; STROUD, Christopher (Org.). *Panorama do Português oral de Maputo*. v. III; *Estruturas gramaticais do Português: problemas e exercícios*. Maputo: Instituto Nacional do Desenvolvimento da Educação, 1998. p. 59-87 (Cadernos de pesquisa do Inde n. 27).

GONÇALVES, Rita. *A Colocação dos Pronomes Clíticos no Português Oral de S. Tomé: Análise e discussão de contextos numa perspectiva comparativa*. 2009. Dissertação (Mestrado) – Universidade de Lisboa, Lisboa, 2009.

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

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

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

LOBO, Tânia. *A colocação dos clíticos em Português: duas sincronias em confronto*. 1992. Dissertação (Mestrado) – Universidade de Lisboa, Lisboa, 1992.

LUCCHESI, Dante; BAXTER, Alan; RIBEIRO, Ilza. *O português afro-brasileiro*. Salvador: EDUFBA, 2009.

MARTINS, Ana Maria. *Clíticos na história do Português*. 1994. Tese (Doutorado) – Universidade de Lisboa, Lisboa, 1994.

MARTINS, Ana Maria. Posição dos pronomes pessoais clíticos. In: RAPOSO, Eduardo; NASCIMENTO, Maria Fernanda; MOTA, Maria Antónia; SEGURA, Luísa; MEN-

DES, Amália (Org.). *Gramática do Português*. Lisboa: Fundação Calouste Gulbenkian, 2013. p. 2231-2302.

MARTINS, Marco Antônio. *Competição de gramáticas do português na escrita catari-nense dos séculos 19 e 20*. 2009. Tese (Doutorado) – Universidade Federal de Santa Catarina, Florianópolis, 2009.

NUNES, Carla da Silva. *Um estudo sobre a ordem dos clíticos em complexos verbais no PB e no PE*. 2009. Dissertação (Mestrado em Letras Vernáculas), Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2009.

NUNES, Carla da Silva. *A ordenação do clítico “se” em complexos verbais nas produções escritas do Brasil e de Portugal nos séculos XIX e XX segundo a perspectiva socio-linguística*. 2014. Tese (Doutorado em Letras Vernáculas) – Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2014.

NUNES, Jairo. Direção de cliticização, objeto nulo e pronome tônico na posição de objeto em Português Brasileiro. In: ROBERTS, Ian; KATO, Mary (Org.). *Português Brasileiro: uma viagem diacrônica*. Campinas: Editora da UNICAMP, 1993. p. 207-222.

PAGOTTO, Emílio Gozze. *A posição dos clíticos em Português: um estudo diacrônico*. 1992. Dissertação (Mestrado), UNICAMP, Campinas, SP, 1992.

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

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

SEMEDO, Manuel Brito. *A colocação dos clíticos no Português em Maputo*. Maputo: Instituto Nacional do Desenvolvimento da Educação, 1997. (Cadernos de pesquisa do Inde no 25)

VIEIRA, Maria de Fátima. *A ordem dos clíticos pronominais nas variedades urbanas europeia, brasileira e são-tomense: uma análise Sociolinguística do Português no início do século XXI*. 2016. Tese (Doutorado em Língua Portuguesa) – Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2016.

VIEIRA, Sílvia Rodrigues. *Colocação pronominal nas variedades europeia, brasileira e moçambicana: para a definição da natureza do clítico em Português*. 2002. Tese (Doutorado em Língua Portuguesa) – Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2002.

VIEIRA, Sílvia Rodrigues. Entre o variável e o categórico: a concordância verbal e a colocação pronominal em variedades do Português. In: REZENDE, Letícia; SILVA, Odair Luiz; MENDONÇA, Marina Célia; ZAVAGLIA, Claudia; BRUNELLI, Anna Flora (Org.). *A interdisciplinaridade e a especificidade linguística: teorias e práticas*. Araraquara; FCL-UNESP Laboratório Editorial; São Paulo: Cultura Acadêmica, 2014. p. 75-98.

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

VIEIRA, Silvia Rodrigues; CORRÊA, Cristina Márcia Monteiro de Lima. Colocação pronominal no Português do Brasil: a contribuição de estudos de percepção auditiva. *Letras de Hoje*, Porto Alegre, v. 52 (1), p. 87-96, 2017.

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

# A ORGANIZADORA

## **SILVIA FIGUEIREDO BRANDÃO**

Professora Titular de Língua Portuguesa (UFRJ), Doutora em Letras Vernáculas (1988, UFRJ), atua, desde 1978, na Área de Língua Portuguesa, no Departamento de Letras Vernáculas da Faculdade de Letras da UFRJ. Em 2009, realizou estágio pós-doutoral, com bolsa CAPES, no Centro de Linguística da Universidade de Lisboa. É bolsista de Produtividade em Pesquisa do CNPq desde 2008 e Cientista do Nosso Estado pela FAPERJ. No Programa de Pós-graduação em Letras Vernáculas, desde 1990, desenvolve e orienta pesquisas no âmbito da Sociolinguística Variacionista e da Dialectologia, em especial nos campos da Fonética-Fonologia e da Morfossintaxe. Entre 1994 e 1996, coordenou o GT de Sociolinguística da ANPOLL, e, entre 2000 e 2004, o Projeto VARPORT (Análise Contrastiva de Variedades do Português), de Cooperação Internacional – CAPES-FCT (Brasil-Portugal). Em 2014, foi eleita Vogal da ALFAL, associação em que também coordena o Projeto 21. Autora de *A geografia linguística no Brasil* (São Paulo: Ática, 1991), com publicações no Brasil e no exterior, organizou, com Maria Antônia Mota, da Universidade de Lisboa, *Análise contrastiva de variedades do Português: primeiros estudos* (Rio de Janeiro: In-Fólio, 2003), e, com Silvia Rodrigues Vieira, *Morfossintaxe e ensino de Português: reflexões e propostas* (Faculdade de Letras-UFRJ, 2004) e *Ensino de gramática: descrição e uso* (São Paulo: Contexto, 2007).



# O PREFACIADOR

## **TJERK HAGEMEIJER**

Professor auxiliar do Departamento de Linguística Geral e Românica da Faculdade de Letras da Universidade de Lisboa e investigador do Centro de Linguística da Universidade de Lisboa (CLUL). Tem como principal área de investigação a sintaxe descritiva e teórica e a formação e evolução diacrónica dos crioulos do Golfo da Guiné, tendo também desenvolvido trabalho sobre a gramática e o contexto sociolinguístico do português falado em África, com especial enfoque no contexto linguístico de São Tomé e Príncipe. Foi coordenador do projeto “As Origens e o Desenvolvimento de Sociedades Crioulas no Golfo da Guiné: Um Estudo Interdisciplinar” (2011-2014), financiado pela Fundação para a Ciência e Tecnologia (FCT), e colaborador do *Atlas of Pidgin and Creole Structures* (Oxford: OUP, 2013), uma iniciativa do MPI-Leipzig (2006-2013). Desde 2008, coordena o projeto Variedades Africanas do Português (VAPOR) do CLUL, com Amália Mendes. Tem também colaborado em iniciativas de defesa e normalização dos crioulos de São Tomé e Príncipe, sendo coautor do Dicionário Livre do Santome-Português (São Paulo: Hedra, 2013) e colaborador da criação de uma proposta de escrita destas línguas, oficializada em 2013. Foi vogal da direção do CLUL (2008-2013) e investigador responsável do grupo de Investigação Anagrama do CLUL (2014-2016).



# AS COLABORADORAS

## **ALESSANDRA DE PAULA**

Professora Adjunta de Língua Portuguesa e Filologia Românica (UERJ), atua no Departamento de Letras da Faculdade de Formação de Professores da Universidade do Estado do Rio de Janeiro. Doutora em Letras Vernáculas (2015, UFRJ), foi bolsista do CNPq e realizou Estágio de Doutorado Sanduíche na Universidade de Lisboa com apoio da CAPES. É pesquisadora da FAPERJ e dedica-se à investigação da mudança linguística no sistema fonológico do português, com ênfase na fala fluminense. Tem artigos publicados, em especial, sobre o vocalismo átono do Português do Brasil.

## **DANIELLE KELY GOMES**

Professora Adjunta II do Departamento de Letras Vernáculas da Universidade Federal do Rio de Janeiro, é Doutora e Mestre em Letras Vernáculas pela Universidade Federal do Rio de Janeiro (2012 e 2006, respectivamente), Bacharel e Licenciada em Português/Literaturas pela Universidade Federal do Rio de Janeiro (2003/2004). Foi professora docente I nas Secretarias Municipal e Estadual de Educação do RJ (2005-2013). Entre 2013 e 2014, foi professora Adjunta I do Departamento de Letras Clássicas e Vernáculas da Universidade Federal Fluminense. É uma das organizadoras de *Uma história de investigações sobre a língua portuguesa* (Blucher, 2018). Dedicar-se a pesquisas em teorias fonológicas, nas interfaces aquisição da linguagem/aprendizagem da escrita, contato linguístico e análises contrastivas entre variedades do Português.

## **FABIANE DE MELLO VIANNA DA ROCHA TEIXEIRA RODRIGUES DO NASCIMENTO**

Doutora (2018) e Mestre (2013) em Letras Vernáculas pela UFRJ e Graduada em Letras/Latim (2011), também pela UFRJ. Atualmente atua como Professora Substituta de Língua Portuguesa na UFRJ, onde exerceu o mesmo cargo também no período de 2012 a 2014. Suas pesquisas vinculam-se à linha “Língua e sociedade: variação e mudança”. No Mestrado, integrou o projeto *O falar fluminense: perfil geo-sociolinguístico – o vocalismo pretônico e desenvolveu, com bolsa do CNPq, pesquisa sociolinguística e difusionista acerca das vogais médias pretônicas na fala de Nova Iguaçu. Já no Doutorado, integrou o projeto *O vocalismo átono no Português do Brasil e no Português de São Tomé* e analisou, com bolsas do CNPq e Nota 10 da FAPERJ, o mesmo fenômeno variável nesta última variedade. Com bolsa CAPES, realizou estágio de Doutorado Sanduíche (2017), no Centro de Linguística da Universidade de Lisboa. De 2008 a 2010, foi monitora da disciplina Fonologia do Português, no Departamento de Letras Vernáculas da Faculdade de Letras da UFRJ. Em sua trajetória acadêmica, tem participado de eventos científicos e publicado artigos de cunho variacionista.*

## **KAREN CRISTINA DA SILVA PISSURNO**

Doutoranda em Letras Vernáculas (UFRJ). Entre 2010 e 2012, atuou como bolsista (CNPq) no projeto “Estudo comparado dos padrões de concordância em variedades africanas, brasileiras e europeias do Português” (ALFAL), coordenado pelas doutoras Silvia Rodrigues Vieira e Silvia Figueiredo Brandão. Na ocasião, realizou trabalhos de coleta, codificação e análise de dados, especialmente no que tangia à Concordância Verbal, do Português de São Tomé e

Príncipe. Em 2015, iniciou a pesquisa de mestrado sobre a variedade de Português falada em Moçambique, pouco explorada até então, o que resultou em uma pesquisa de campo, realizada em 2016, juntamente com Silvia Rodrigues Vieira, para a elaboração de um *corpus* linguístico da variedade em questão, através da execução de entrevistas sociolinguísticas em Maputo, Moçambique. Os resultados da investigação foram apresentados na dissertação de mestrado, defendida em 2017. Em 2018, a pesquisa de doutorado permanece no campo das variedades africanas, com ênfase na moçambicana, com o intuito de continuar desenvolvendo a Fase 2 do Projeto ALFAL (*Estudo comparado de padrões de concordância em variedades africanas, brasileiras e europeias do Português: a natureza das restrições e o contato linguístico*), dando atenção especial à situação de multilinguismo, fortemente observada na variedade em estudo.

#### MARIA DE FATIMA VIEIRA

Doutora (2016) e Mestra (2011) em Letras Vernáculas na área de Língua Portuguesa pela Universidade Federal do Rio de Janeiro (UFRJ). Possui Graduação e Licenciatura em Português-Espanhol (2008) pela mesma instituição. Atualmente, trabalha como Docente II da Prefeitura Municipal de Angra dos Reis e como Professor I na Prefeitura Municipal do Rio de Janeiro. Tem experiência na área de Letras, com ênfase em Língua Portuguesa, atuando principalmente nos seguintes temas: Sociolinguística Variacionista; Variedades do Português; Cliticização pronominal. Sua produção acadêmica conta com diversos trabalhos apresentados em congressos, resumos, além da publicação de artigos e do livro intitulado *A colocação pronominal no Português Europeu oral contemporâneo – A cliticização pronominal em lexias verbais simples e em complexos verbais: uma investigação sociolinguística* (Novas Edições Acadêmicas – Saarbrücken, 2014) – obra que sintetiza sua pesquisa no âmbito do referido Mestrado.

#### RAPHAELA RIBEIRO PASSOS

Mestre em Letras Vernáculas, Língua Portuguesa, pela Universidade Federal do Rio de Janeiro (2018) e Graduada com licenciatura plena em Letras, habilitação em Português e Espanhol (2015) pela mesma universidade. Durante a graduação, atuou como Monitora de Literatura Espanhola e como Monitora de Espanhol do CLAC – Cursos de Línguas Abertos à Comunidade –, além de ter realizado Iniciação Científica, com Bolsa PIBIC/UFRJ, na área de Fonética e Fonologia. No Mestrado, dedicou-se à pesquisa vinculada ao sistema fonológico do Português de São Tomé e Príncipe.

#### SILVIA RODRIGUES VIEIRA

Professora Associada de Língua Portuguesa (UFRJ), Doutora em Letras Vernáculas (2002, UFRJ), atua desde 1997, na Área de Língua Portuguesa, no Departamento de Letras Vernáculas da Faculdade de Letras da UFRJ. É bolsista de Produtividade em Pesquisa do CNPq desde 2015 e Cientista do Nosso Estado pela FAPERJ. No Programa de Pós-graduação em Letras Vernáculas, desde 2003, desenvolve e orienta pesquisas no âmbito da Sociolinguística Variacionista, em especial no campo da Morfossintaxe. Igualmente no âmbito desse Programa, na linha Língua e Ensino, e no Mestrado Profissional em Letras/PROFLETRAS, orienta também trabalhos no campo do Ensino de gramática. Atua no GT de Sociolinguística da ANPOLL coordenando o Eixo Sociolinguística e ensino, e coordena o Projeto ALFAL 21 – *Estudo comparado dos padrões de concordância em variedades africanas, brasileiras e europeias do Português*. Com publicações no Brasil e no exterior, organizou, dentre outras, as seguintes obras: com Silvia Figueiredo Brandão, *Morfossintaxe e ensino de Português: reflexões e propostas* (Fac. Letras-UFRJ, 2004) e *Ensino de gramática: descrição e uso* (Contexto, 2007); individualmente, *A concordância verbal em variedades do Português: a interface Fonética: Morfossintaxe* (Vermelho Marinho, 2015); *Variação e ensino de gramática: diagnose e propostas pedagógicas* (Fac. Letras-UFRJ, 2017).