



**UNIVERSIDADE DE BRASÍLIA
INSTITUTO DE LETRAS
DEPARTAMENTO DE LINGÜÍSTICA, PORTUGUÊS E LÍNGUAS CLÁSSICAS
PROGRAMA DE PÓS-GRADUAÇÃO EM LINGÜÍSTICA**

TESE DE DOUTORADO

**PORTUGUÊS DE SÃO TOMÉ (E PRÍNCIPE): LEXICALIZAÇÃO,
GRAMATICALIZAÇÃO E *CODE-SWITCHING* EM CONTEXTO MULTILÍNGUE
*SÃO-TOMÉAN PORTUGUESE: LEXICALIZATION, GRAMMATICALIZATION AND
CODE-SWITCHING IN A MULTILINGUAL SETTING***

JULWAIY QUARESMA CARDOSO PIMENTEL NETO

**BRASÍLIA
2025**



**UNIVERSIDADE DE BRASÍLIA
INSTITUTO DE LETRAS
DEPARTAMENTO DE LINGUÍSTICA, PORTUGUÊS E LÍNGUAS CLÁSSICAS
PROGRAMA DE PÓS-GRADUAÇÃO EM LINGUÍSTICA**

JULWAITY QUARESMA CARDOSO PIMENTEL NETO

**PORTUGUÊS DE SÃO TOMÉ (E PRÍNCIPE): LEXICALIZAÇÃO,
GRAMATICALIZAÇÃO E *CODE-SWITCHING* EM CONTEXTO MULTILÍNGUE**

Tese de doutorado de minha autoria exclusiva apresentada ao Programa de Pós-Graduação em Linguística do Departamento de Linguística, Português e Línguas Clássicas do Instituto de Letras da Universidade de Brasília como requisito parcial para obtenção do grau de Doutor em Linguística. Área de concentração: Teoria e Análise Linguística. Linha de pesquisa: Léxico e Terminologia.

Orientadora: Profa. Dra. Enilde Faulstich

**BRASÍLIA
2025**

FOLHA CATALOGRÁFICA

Qp

QUARESMA CARDOSO PIMENTEL NETO, Julwaity
Português de São Tomé (e Príncipe): Lexicalização,
Gramaticalização e Code-switching em contexto multilíngue /
Julwaity QUARESMA CARDOSO PIMENTEL NETO; orientador Enilde
Leite de Jesus Faulstich. Brasília, 2025.
165 p.

Tese(Doutorado em Linguística) Universidade de Brasília,
2025.

1. Lexicalização. 2. Gramaticalização. 3. Code-switching.
4. Contato de línguas. 5. Português de São Tomé. I. Leite de
Jesus Faulstich, Enilde, orient. II. Título.

BANCA EXAMINADORA

PORTUGUÊS DE SÃO TOMÉ (E PRÍNCIPE): LEXICALIZAÇÃO, GRAMATICALIZAÇÃO E *CODE-SWITCHING* EM CONTEXTO MULTILÍNGUE

Doutorando: Julwaity Quaresma Cardoso Pimentel Neto

COMPOSIÇÃO DA BANCA EXAMINADORA

Profª. Dra. Enilde Leite de Jesus Faulstich UnB/PPGL/LIP/IL	-	Orientadora/Presidente
Prof. Dr. Paulo Medeiros Júnior UnB/PPGL/LIP/IL	-	Membro efetivo interno
Profª. Dra. Rosa Amélia Pereira da Silva IFB/PREN (Pró-Reitoria de Ensino)	-	Membro efetivo externo
Profª. Dra. Soraia da Rosa Mendes UnICEUB/PPGD (Programa de Pós-Graduação em Direito)	-	Membro efetivo externo
Profª. Dra. Michelle Machado de Oliveira Vilarinho UnB/PPGL/LIP/IL	-	Membro suplente

Brasília-DF, 10 de dezembro de 2025.

DEDICATÓRIA

Dedico esta tese à minha mãe,
Sra. Cucu (“Coração”), *in*
memoriam, por tudo que
representou e representa na
minha vida e na minha
caminhada, e ao meu primo
Alexandre Cardoso (“Michel
Pato”), *in memoriam*, que me
apelidou de Doutor quando eu
ainda estava no ensino
fundamental em São Tomé.

AGRADECIMENTOS

À querida Professora Doutora Enilde Faulstich, pela sugestão do tema, pelas cobranças firmes, paciência e ensinamentos.

À minha mãe, pelo incentivo e esclarecimentos de muitas expressões do Santome.

À minha esposa, Nida Marques, pelo amor e pela paciência.

Aos professores do Programa de Pós-Graduação em Linguística – PPGL/UnB.

À Profa. Dra. Ulisdete Rodrigues pelo apoio e incentivo.

Ao Instituto Federal de Brasília (IFB), pelas licenças para capacitação concedidas.

Aos(Às) colegas do IFB Profa. Dra. Veruska Machado, Profa. Dra. Maria Aparecida e Prof. Me. Lucas Barbosa, pelo apoio.

Aos amigos Santolas Leovegildo da Mata, Nascimento Rosamonte, Armando Monteiro, Joimilte Bonfim, N'Diê Nobre, Dekker Batista, Belmiro Soares, Jessica Monteiro e Diógenes Aguiar, pela força e contribuição durante a pesquisa.

A todos que contribuíram para o desenvolvimento desta pesquisa, meus sinceros agradecimentos.

EPÍGRAFE

*Piskadô ku toma kanwa p'ê ba omali
sa pund'ê tê fôsa mat'unwa
ngandu.ST*

O pescador que pega uma canoa e
vai ao mar tem força para pescar um
tubarão.^{PT}

Provérbio são-tomense

Sentido:

Se eu me propus a fazer essa
pesquisa, é porque sei que sou
capaz.

RESUMO

Esta tese visa analisar os processos de lexicalização e gramaticalização no Português de São Tomé e o fenómeno de *code-switching* português-santome. São Tomé e Príncipe é um cenário de línguas em contato com três crioulos autóctones – Santome (crioulo predominante), Ngola e Lung'le – além do Kabuverdianu, e o Português, a única língua oficial. Por se inserir num contexto multilíngue, esta pesquisa segue a abordagem de lexicalização e gramaticalização discutida por Meillet (2020 [1912]), Moreno Cabrera (1998), Castilho (2003), Neves (2004, 2021), Gonçalves et al. (2007), Bybee (2016, 2020), entre outros. Para discussão sobre línguas em contato e *code-switching*, esta tese se ampara nas perspectivas de Grosjean (1982), Gumperz (1982), Tarallo & Alkmin (1987), Calvet (2002) e Mozzillo (2009). A pesquisa seguiu uma abordagem qualitativa com o propósito de descrever itens lexicais e estruturas sintáticas do Português de São Tomé assim como o uso do *code-switching* (Português-Santome) nas interações sociais entre os falantes. O corpus utilizado no presente trabalho de pesquisa é composto de dados coletados em textos multimodais, entre os quais uma crônica, artigos de jornais digitais locais, livros de contos tradicionais e história de São Tomé e Príncipe e, principalmente, publicações em mídias sociais, sem recorte temporal. A análise morfológica desses itens lexicais, suas estruturas sintáticas e os processos de lexicalização e gramaticalização do Português de São Tomé demonstraram que grande parte do léxico e construções sintáticas do Português de São Tomé são empréstimos, decalques e/ou derivações do Santome, motivados pelo contato linguístico, indicando um impacto importante do Santome no Português de São Tomé. Além disso, o recurso frequente ao *code-switching* Português-Santome pelos falantes nessa comunidade de fala se dá primordialmente por questões de identidade e pertencimento e pela necessidade de expressão do pensamento que o Português de São Tomé, sozinho, não é capaz de fazer na sua plenitude. Portanto, a compreensão e análise do português falado em São Tomé e Príncipe dependem em grande medida do conhecimento e domínio da principal língua nativa de São Tomé e Príncipe, o Santome.

Palavras-chave: Lexicalização; Gramaticalização; Contexto Multilíngue; Línguas em Contato; Português de São Tomé, Santome; *Code-Switching*.

ABSTRACT

This thesis aims to analyze the lexicalization and grammaticalization processes in São Tomé Portuguese and code-switching Portuguese-Santome. São Tomé and Príncipe is a setting of language contact, with three indigenous creoles – Santome (the main creole), Ngola, and Lung'le – besides Kabuverdianu, and Portuguese, the only official language. As a multilingual setting, this research follows the lexicalization and grammaticalization approaches discussed by Meillet (2020 [1912]), Moreno Cabrera (1998), Castilho (2003), Neves (2004, 2021), Gonçalves et al. (2007), Bybee (2016, 2020), among others. To discuss language contact and code-switching, this thesis is based on the perspectives of Wong (1979), Grosjean (1982), Gumperz (1982), Tarallo & Alkmin (1987), Calvet (2002) and Mozzillo (2009). The research followed a qualitative approach with the purpose of analyzing and describing lexical items and syntactic structures of São Tomé Portuguese as well as the use of code-switching (Portuguese-Santome) in social interactions between speakers. The corpus used in this research is composed of data collected from multimodal texts, including a chronicle, articles from local online newspapers, books of traditional stories and the history of São Tomé and Príncipe, and, primarily, social media posts, without a time frame. Morphological analysis of these lexical items, their syntactic structures, and the lexicalization and grammaticalization processes of São Toméan Portuguese demonstrated that a large portion of the lexicon and syntactic constructions of São Toméan Portuguese are loanwords, tracings, and/or derivations of Santome, motivated by language contact, indicating a significant impact of Santome on São Toméan Portuguese. Furthermore, the frequent use of code-switching Portuguese-Santome by speakers in this speech community occurs primarily for reasons of identity and belonging and the need to express thought that São Toméan Portuguese alone cannot fully achieve. Therefore, the understanding and analysis of Portuguese spoken in São Tomé and Príncipe depend largely on the knowledge and mastery of the main native language of São Tomé and Príncipe, Santome.

Keywords: Lexicalization; Grammaticalization; Multilingual Setting; Language Contact Sao-Tomean Portuguese; Santome Creole; Code-Switching.

LISTA DE FIGURAS

Figura 1 - Ba kwê.....	22
Figura 2 - R DSTP no mundo e no Golfo da Guiné.....	26
Figura 3 - Forte de São Sebastião/ Estátuas dos navegadores portugueses.....	27
Figura 4 - Estátua de Rei Amador na Cidade de São Tomé.....	32
Figura 5 - São Tomé: posição geográfica no Golfo da Guiné, topografia (A), demografia (B).....	35
Figura 6 - Bandeira de São Tomé e Príncipe.....	36
Figura 7 - Braço de Armas de São Tomé e Príncipe.....	37
Figura 8 - Nota de Dbs. 50.000.....	39
Figura 9 - Vianteiro.....	42
Figura 10 - Mapa político de STP.....	44
Figura 11 - Baía de Ana Chaves / Cidade de São Tomé.....	45
Figura 12 - Slogan do partido MLSTP nas Eleições legislativas de 2022.....	60
Figura 13 - Multissistema Linguístico de Castilho.....	68
Figura 14 - Tipos de code-switching.....	87
Figura 15 - Code-switching intersentencial.....	88
Figura 16 - Postagem de CS.....	94
Figura 17 – Fulumento.....	105
Figura 18 - Wexaismo: processo de formação.....	107
Figura 19 - Olho cheio e suas variações.....	108
Figura 20 - Code-switching intersentencial entre falantes/usuários diferentes.....	111
Figura 21 - Ontá ocê bom (Vava Neviense, cantor são-tomense).....	113
Figura 22 - Hotel Omali São Tomé.....	119
Figura 23 - Verbo “yonar” e expressões “abrir fundá” e “tomar parte”.....	128
Figura 24 - Nova Fatura Vêndê Pleçu Dóxi	130
Figura 25 - Postagem de JB.....	131
Figura 26 - Processos linguísticos na Postagem de JB.....	133
Figura 27 - Postagem de PQ: Code-switching	134
Figura 28 - Postagem do GS: Code-switching	136
Figura 29 - Advérbio “já” no PST e Ideofones “é” e “ô”	143

LISTA DE QUADROS

Quadro 1 - Distribuição da população são-tomense por distrito e faixa etária: RGPH-2012.....	18
Quadro 2 - Evolução da população são-tomense por faixa etária de 2001 a 2012.....	20
Quadro 3 - População residente por distrito de residência atual por local de nascimento	46
Quadro 4 - Vogais orais breves.....	53
Quadro 5 - Vogais nasais / nasalidade precedida de /p/ e /b/.....	54
Quadro 6 - Vogais nasais - nasalidade precedida de grafemas diferentes de /p/ e /b/.....	54
Quadro 7 - Consoantes segundo ALUSTP.....	55
Quadro 8 - Partidos políticos na língua Santome.....	59
Quadro 9 - Paradigma Formal vs Paradigma Funcional.....	63
Quadro 10 - Matéria sobre Ontá ocê bom.....	115
Quadro 11 - Fauna.....	117
Quadro 12 - Flora.....	118
Quadro 13 - Topônimos	118
Quadro 14 - Projetos sociais e ambientais	120
Quadro 15 - Jornais digitais e/ou páginas nas mídias digitais.....	120
Quadro 16 - Ideofones do PST/ST.....	149

LISTA DE GRÁFICOS

Gráfico 1 – Proporção da população por línguas faladas a nível geral.....	52
---	----

LISTA DE ABREVIATURAS E SIGLAS

1PL	PRIMEIRA PESSOA DO PLURAL
1SG	PRIMEIRA PESSOA DO SINGULAR
2PL	SEGUNDA PESSOA DO PLURAL
2SG	SEGUNDA PESSOA DO SINGULAR
3PL	TERCEIRA PESSOA DO PLURAL
3SG	TERCEIRA PESSOA DO SINGULAR
ADJ	ADJETIVO
ADV	ADVÉRPIO
ALUSTP	ALFABETO UNIFICADO DAS LÍNGUAS NATIVAS DE SÃO TOMÉ E
PRÍNCIPE	
ANAT	ANATOMIA
ART	ARTIGO
ASP	ASPECTO
BOT	BOTÂNICA
CF	CONFIRA, CONFRONTE, COMPARE
CONJ	CONJUNÇÃO
CTR	CONTRAÇÃO
CS	CODE-SWITCHING
DEM	PRONOME DEMONSTRATIVO
EX:	EXEMPLO
GRA	GRAMATIZALIZAÇÃO
FAU	FAUNA
FLO	FLORA
IDEO	IDEOFONE
INT	INTERJEIÇÃO
L1	LÍNGUA MATERNA
L2	SEGUNDA LÍNGUA
LEX	LEXICALIZAÇÃO
LIT	LITERAL
MET	METAFÓRICO

MLSTP	MOVIMENTO DE LIBERTAÇÃO DE SÃO TOMÉ E PRÍNCIPE
INE	INSTITUTO NACIONAL DE ESTATÍSTICA DE SÃO TOMÉ E PRÍNCIPE
OBJ	OBJETO
NOM	NOME
N.P.	NÃO PAGINADO
P	PÁGINA
PB	PORTUGUÊS DO BRASIL
PL	PLURAL
POS	POSSESSIVO
PP	PORTUGUÊS DE PRÍNCIPE
PPT	PORTUGUÊS DE PORTUGAL
PRED	PREDICADO
PREP	PREPOSIÇÃO
PRON	PRONOME
PST	PORTUGUÊS DE SÃO TOMÉ
PT	PORTUGUÊS
RGPH	RECENSEAMENTO GERAL DA POPULAÇÃO E HABITAÇÃO
SIN	SINÔNIMO
ST	SANTOME
STP	SÃO TOMÉ E PRÍNCIPE
SUB	SUBSTANTIVO
SUJ	SUJEITO
VAR	VARIANTE
VERB	VERBO

SUMÁRIO

INTRODUÇÃO	16
CAPÍTULO 1: SÃO TOMÉ E PRÍNCIPE	25
1.1. CONTEXTUALIZAÇÃO HISTÓRICA.....	25
1.1.1. Os Angolares.....	30
1.1.2. Símbolos Nacionais de São Tomé e Príncipe.....	36
1.2. CONTEXTO ATUAL.....	43
1.3. LÍNGUAS.....	48
1.3.1. Línguas Nativas de São Tomé e Príncipe.....	48
1.3.2. Alfabeto Unificado das Línguas Nativas de São Tomé e Príncipe (ALUSTP) ..	52
1.3.3. Língua Portuguesa em São Tomé e Príncipe	56
1.4 SÍNTESE DO CAPÍTULO	60
CAPÍTULO 2: PERSPECTIVAS TEÓRICAS	62
2.1. LEXICALIZAÇÃO.....	69
2.2. GRAMATICALIZAÇÃO.....	74
2.3. CODE-SWITCHING.....	84
2.4. A TESE.....	89
CAPÍTULO 3: METODOLOGIA DE PESQUISA	90
3.1. PROCEDIMENTOS METODOLÓGICOS.....	90
3.2 A PESQUISA: TIPO E NATUREZA.....	91
3.3. LEVANTAMENTO BIBLIOGRÁFICO	91
3.4. DADOS: DELIMITAÇÃO E COLETA.....	97
3.5 SÍNTESE DO CAPÍTULO	99
CAPÍTULO 4: ANÁLISE DOS DADOS	100
4.1 PROCEDIMENTOS DE ANÁLISE.....	100
4.2 LEXICALIZAÇÃO.....	102
4.2.1 Análise de itens lexicais sob o ponto de vista da morfologia.....	102

4.2.2. Nomes comuns emprestados do Santome e sem alteração.....	102
4.2.3. Itens lexicais resultantes de processos parciais de mudança linguística....	103
4.2.4. Itens lexicais resultantes de processos de mudança linguística por derivação.....	104
4.2.5. Itens lexicais emprestados do Santome: fauna, flora, topônimos e projetos	117
4.3. GRAMATICALIZAÇÃO.....	120
4.4. CODE-SWITCHING.....	131
4.4.1. Provérbios ou ditados populares no PST.....	138
4.4.2. Ideofones no Português de São Tomé	147
4.5 SÍNTESE DO CAPÍTULO	153
 CONSIDERAÇÕES FINAIS.....	 154
REFERÊNCIAS.....	158

INTRODUÇÃO

Esta pesquisa se insere num contexto de discussão sobre questões linguísticas relativas a São Tomé e Príncipe. Por sua vez, os tópicos abordados surgem da identificação de uma lacuna não preenchida até o atual estado da arte. O título desta tese – Português de São Tomé (e Príncipe): lexicalização, gramaticalização e *code-switching* em contexto multilíngue – é fruto, justamente, de uma inquietação muito antiga e considerada muito relevante para os estudos linguísticos em geral e para o caso de São Tomé (e Príncipe) em particular.

“E Príncipe” é usado intencional e geralmente entre parênteses quando se referir ao estudo porque, embora o país seja territorial e populacionalmente muito pequeno, as duas ilhas situam-se a 150 km de distância, no Oceano Atlântico. O Lung’le (ou Principense) é falado praticamente apenas na Ilha do Príncipe, além do Santome e Kabuverdianu. No entanto, “e Príncipe” é mantido por uma questão de identidade e unidade nacional e por entendermos que os processos e fenômenos aqui abordados também ocorram na Ilha do Príncipe. Nesse sentido, nesta pesquisa, utiliza-se Português de São Tomé (doravante PST) e Português do Príncipe (PP) separadamente, pois existem diversos trabalhos de pesquisa relacionados ao PST e outros sobre o PP. Ainda assim, a pergunta costumeira e clássica que ouvimos com frequência quinzenal ou mensal como morador de Brasília-DF é “Em São Tomé, vocês falam Português de Portugal, né?”

Foram avaliados diversos trabalhos durante a pesquisa, mas a maioria geralmente aborda questões específicas do cenário linguístico em São Tomé e Príncipe como fonética e fonologia, léxico, desafios do ensino do português como L1 entre outros.

Porém, nenhum dos trabalhos se dedica a questões mais abrangentes, nem discute objetiva e especificamente como ocorrem os processos de lexicalização e de gramaticalização do PST que são também muito importantes nesse ambiente multiétnico e multilíngue em que o PST está inserido, que é, na verdade, um produto em constante evolução e mudança causada por contatos, principalmente pelo Santome, que é a língua crioula dominante, muito menos o *code-switching*. O *code-switching*, também chamado de mudança de código ou alternância de línguas, é um fenômeno muito recorrente na vida prática de comunicação e interação social entre os são-tomenses.

Além disso, o trabalho se faz necessário e mais relevante ainda quando realizado a partir da perspectiva de nativo de São Tomé e Príncipe e falante de Santome. O Santome possui níveis ou registros de linguagem com características bem particulares: a língua é usada com muito maior frequência e fluência em determinadas regiões ou comunidades, principalmente pesqueiras, como Praia Gamboa, Praia Cruz, Micoló, Guadalupe, Praia de Morro Peixe, Praia Melão, Pantufo, Neves assim como em bairros da periferia da capital São Tomé como, Ponte Graça, Mulundu e São João da Vargem. Portanto, fato é que o nível ou domínio do Santome e do Ngola, nessas regiões, é muito maior que nas demais regiões.

O Santome é uma língua metafórica e se utiliza de provérbios com grande frequência. Portanto, a compreensão da língua, visão do mundo e expressão do pensamento dependem, em grande medida, dos saberes populares e das referências comuns entre os interlocutores. Os provérbios e as metáforas são popularmente conhecidos, em São Tomé e Príncipe, como “verso” e quem mais usa e domina esses

versos são as pessoas mais velhas¹. O Quadro 1 (p. 18) apresenta a distribuição da população são-tomense por distrito e faixa etária em todo o território nacional. Verificamos que, com exceção da faixa etária de zero a cinco anos na Região Autónoma do Príncipe (RAP), todas as faixas etárias da população estão distribuídas, proporcionalmente, de forma bastante uniforme por todos os distritos.

**Quadro 1 – Distribuição da população são-tomense por distrito e faixa etária:
RGPH-2012**

Distritos		Faixa Etária					
		Total	0 – 5	0 – 17	6 – 11	15 – 64	65+
STP	Efetivo	178739 6590	33325	86060	28966	97530	
	%	... 3,7	18,6	48,1	16,2	54,6	
Água Grande	Efetivo	69454 2416	12514	31889	10569	39418	
	%	... 3,5	18,0	45,9	15,2	56,8	
Mé-Zochi	Efetivo	44752 1820	8187	21779	7441	24143	
	%	... 4,1	18,3	48,7	16,6	53,9	
Lobata	Efetivo	19365 709	3863	9514	3170	10331	
	%	... 3,7	19,9	49,1	16,4	53,3	
Lembá	Efetivo	14652 492	2991	7534	2496	7568	

¹ Velha(o) nesse contexto é utilizado no sentido de idoso e sem etarismo. Aliás, de forma geral, em São Tomé a idade é tão importante que a hierarquia começa justamente pela idade. Prova disso é que, embora a famosa frase “Sabe com quem está falando?” também exista no PST (Sabe quem eu sou?), a frase mais comum era (ou é) “Sabe qual é a minha idade?”, que traduzindo seria “Sabe quanto respeito você me deve?”. E essa frase era geralmente seguida, no Santome, por outra - *Un sa maxi tamé men/pe bô*. [Sou mais velho(a) que o(a) seu(sua) pai/mãe], *Un sa maxi tamé dona/dono bô*. [Sou mais velho(a) que o(a) seu(sua) avô/avó] etc. Porém, a referência aos pais ou avós é para dizer o seguinte: Se os seus pais (ou avós) me devem respeito, imagine você.

	%	...	20,4	51,4	17,0	51,7
		3,4				
Cantagalo	Efetivo	17161	3277	8647	2983	9047
		632				
	%	...	19,1	50,4	17,4	52,7
		3,7				
Caué	Efetivo	6031	1097	3069	1034	3150
		246				
	%	...	18,2	50,9	17,1	52,2
		4,1				
R.A.P	Efetivo	7324	241	3628	1273	3888
		275				
	%		3,3	49,5	17,4	53,1
		3,8				

Fonte: Adaptado do INE-STP (2014, p. 65)

O Quadro 2 (p. 20), por sua vez apresenta a evolução da população são-tomense por faixa etária entre os anos de 2001 e 2012. Os dados do quadro mostram que a população idosa são-tomense está em declínio enquanto a população mais jovem está em constante crescimento. Diante desse cenário, podemos afirmar que o ditado popular local “Cada velho que morre é uma biblioteca que queima.” não é um exagero, uma vez que as tradições africanas são baseadas na oralidade, ou seja, a transferência dos saberes e conhecimentos se dá pela tradição oral. Portanto, o número de falantes do Santome está em declínio e, conseqüentemente, as línguas nativas, os costumes, memórias e saberes populares estão também ameaçados, afetando principalmente o Santome de maneira importante.

Quadro 2 – Evolução da população são-tomense por grupo etário de 2001 a 2012

ANO GRUPOS ETÁRIOS	2001		2012	
	Efetivos	%	Efetivos	%
Total	137599	-----	178736	-----
0 – 5	24334	17,68	33325	18,64
0 - 17	22795	16,57	28966	16,21
6 – 11	68389	49,70	86060	48,15
5 - 49	33430	24,30	42893	24,00
5 - 64	73828	53,65	97530	54,57
65+	5897	4,29	6590	3,69

Fonte: INE-STP (2014, p. 65)

O caráter metafórico do Santome se ajusta ao argumento de Bybee (2020, p. 239) quando declara que uma das fontes para novos significados para construções em vias de gramaticalização é a extensão metafórica, pois uma metáfora mapeia uma relação estrutural de um domínio mais concreto para um domínio mais abstrato. A autora afirma ainda que “metáforas especiais baseadas no corpo humano, ou animal, são particularmente frequentes no desenvolvimento de preposições e posposições.

Sobre línguas africanas e a metáfora, Heine et al. (1991b *apud* Bybee, 2020, p. 239) afirmam que:

As línguas africanas são especialmente ricas nesse tipo de desenvolvimento metafórico. Por exemplo, citam o substantivo do suaíli *mbele*, que significa ‘peito’, mas também ‘lado da frente, parte da frente’ quando se aplica a objetos não humanos. A relação espacial entre o peito e o corpo humano todo é retrçada sobre outros objetos por metáfora. *Mbele* também se usa como preposição locativa significando ‘diante de’ e até mesmo como advérbio temporal significando ‘antes’. (Heine et. al., 1991b *apud* Bybee, 2020, p. 239)

No Santome, acontece algo bem semelhante com as palavras *pêtu* [peito] e *wé* [olho]. No Santome, existe a expressão *pêtu-bala – Bô ton pêtu-bala, kwa da ku bô*.ST

[Você se meteu (onde não devia) e teve problemas.]^{PT} No exemplo em questão, *pêtu-bala* tem o sentido de ‘ir adiante’, ‘mostrar esperteza’². Devemos salientar também que *pêtu-bala* foi emprestado do ST e é usado em situações informais no PST e concorre com o item lexical ‘esperteza’, mas apenas tipicamente em posições sintáticas e contextos de uso semelhantes aos do Santome, conforme exemplos a seguir:

(1) *Pêtu-bala*

a) *Bô ton pêtu-bala, kwá da ku bô.*ST

2SG tomar petu-bala coisa dar com 2SG

b) Você tomou *pêtu-bala*, coisa deu com você.^{PST}

2SG tomar petu-bala coisa dar com 2SG

c) Você foi se meter e teve problemas.^{PT}

A palavra *wê*, por sua vez, tem praticamente os mesmos usos e significados de *mbele*. Nos exemplos a seguir, *wê* é usado com sentido de ‘olho’, ‘lado da frente, parte da frente’, ‘progredir’, ‘desenvolver’, ‘ir para frente’ (tanto com conotação positiva como negativa), se aplicando a partes do corpo humano assim como a objetos. À semelhança de *mbele*, *wê* também é usado como preposição locativa significando ‘diante de’, mas não como advérbio temporal significando ‘antes’:

(2) *Bô na tê wê fô?* [Você não tem olhos? / Você não enxerga direito?] – olho (*anat.*);

(3) *Kwa da mu ni wê d’ope?* [Alguma coisa bateu na frente do meu pé.] – canela (lugar, *anat.*);

(4) *Male, fô wê mu?* [Menino, saia da minha frente.] – lugar;

² No PST, ao contrário do adjetivo “esperto” que tem conotação positiva, o substantivo ‘esperteza’ geralmente tem conotação negativa.

(5) *Santome (non) ska ba wê*. [(O nosso) São Tomé (e Príncipe) está progredindo.]

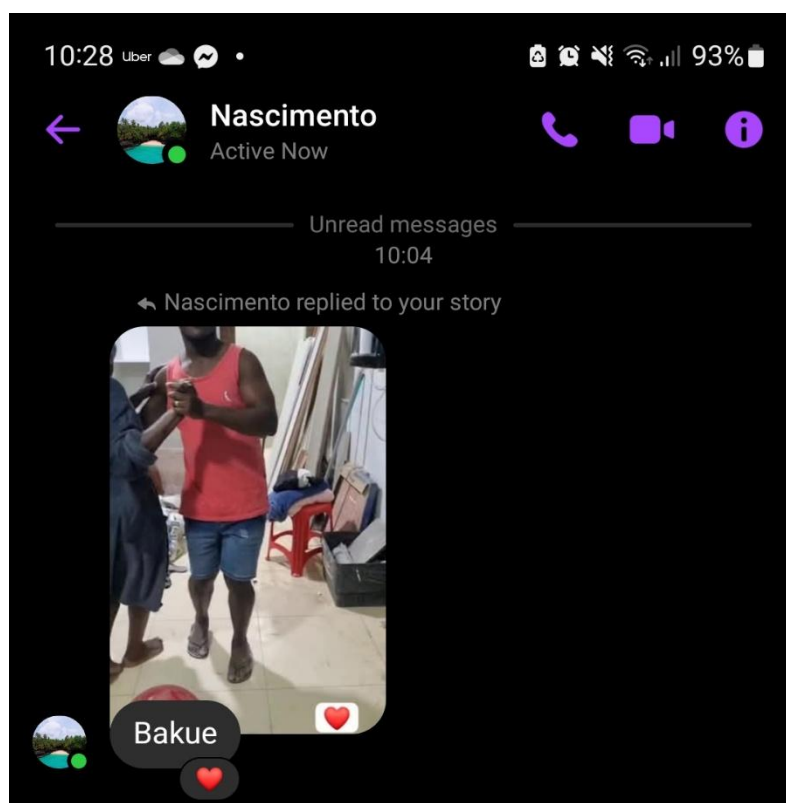
– *ba wê* – ir para frente, progresso;

(6) *Ba kwê! / Ba wê!* [Viva!] – interjeição; (Figura 1)

(7) *Bô pô xa ba wê*. [Vai na esperteza. / Vai se meter.] (conotação negativa)

(8) *Bô fe kwa bô, ami so sa ni wê sode myole*. [Você aprontou e eu que estou diante do policial agora.] - diante de, perante a/o.

Figura 1 – Ba kwê!



Fonte: Autoria própria

O objetivo geral desta tese é estudar o Português de São Tomé (e Príncipe) no seu contexto específico de várias línguas em contato, observando suas características, hierarquias sociais e particularidades no uso quotidiano das línguas. Fazem parte dos

objetivos específicos descrever e analisar os processos de lexicalização e de gramaticalização, contrastar o léxico e a gramática do PST com o léxico e a gramática do Santome, investigar o fenômeno de *code-switching* português-santome e determinar sua motivação e contexto de uso para demonstrar algumas características lexicais e gramaticais singulares que fazem dessa língua uma variante da língua portuguesa – o Português de São Tomé. Para alcançar os objetivos propostos, apresentaremos dados e analisaremos exemplos de processos de lexicalização e de gramaticalização próprios do PST que caracterizam empréstimos de itens lexicais e decalques de estruturas sintáticas do Santome.

Esta tese está organizada em quatro capítulos, além da Introdução e das Considerações Finais, de modo a demonstrar como o processo de colonização de São Tomé e Príncipe e seus desdobramentos resultaram nesse mosaico cultural e linguístico, que é a sociedade são-tomense atual.

A Introdução apresenta o tema, justificativa para a realização do trabalho, os objetivos da pesquisa e a hipótese. Além disso, apresenta a evolução e distribuição demográfica por faixa etária e distritos do país e demonstra como essas dinâmicas sociais impactam o uso das línguas e os processos de mudança linguística no arquipélago.

No primeiro capítulo, “São Tomé e Príncipe”, apresentamos a contextualização histórica do país, desde a chegada dos colonizadores portugueses, os processos e ciclos de povoamento, a exploração econômica por meio dos grandes ciclos com introdução das diferentes culturas/plantações, cana-de-açúcar, cacau e café assim como o contexto atual dessas Ilhas do Equador. É também neste capítulo que discutimos polêmica envolvendo a origem dos Angolares e suas lutas, o processo de descolonização,

independência, a organização política, a atual situação socioeconômica, cultural e, principalmente, o cenário linguístico do país.

No segundo capítulo, “Perspectivas Teóricas”, discutimos lexicalização e gramaticalização em geral, assim como *code-switching*. Apresentamos os conceitos de acordo com os autores que discutem essa temática, a partir da perspectiva funcionalista e estabelecemos a interface com teorias sobre contato de línguas, uma vez que esta pesquisa se concentra num ambiente multilíngue. Nesse sentido, nos apoiamos nos teóricos que se alinham com processos de mudança linguística descritos e exemplos de uso presentes nos dados que sustentam os argumentos defendidos.

No terceiro capítulo, “Metodologia de Pesquisa”, apresentaremos os dados e os caminhos percorridos, assim como as estratégias adotadas para a definição e seleção das fontes, realização e métodos de coleta e tratamento dos dados. Além disso, explicitaremos os procedimentos metodológicos da pesquisa.

No quarto e último capítulo, “Análise dos dados”, descrevemos os processos de lexicalização e gramaticalização em contexto multilíngue, o *code-switching* português-santome e as mudanças linguísticas no PST. É também, neste capítulo, que serão apresentados os exemplos e análises dos dados, embasados em discussões teóricas específicas para, em seguida, apresentar conclusões com base nos objetivos descritos.

Nas Considerações Finais, será apresentada uma revisão geral da pesquisa, os objetivos propostos, desafios enfrentados, os procedimentos adotados e as metas alcançadas. É nesta seção que também serão apresentadas as implicações deste estudo para pesquisas futuras.

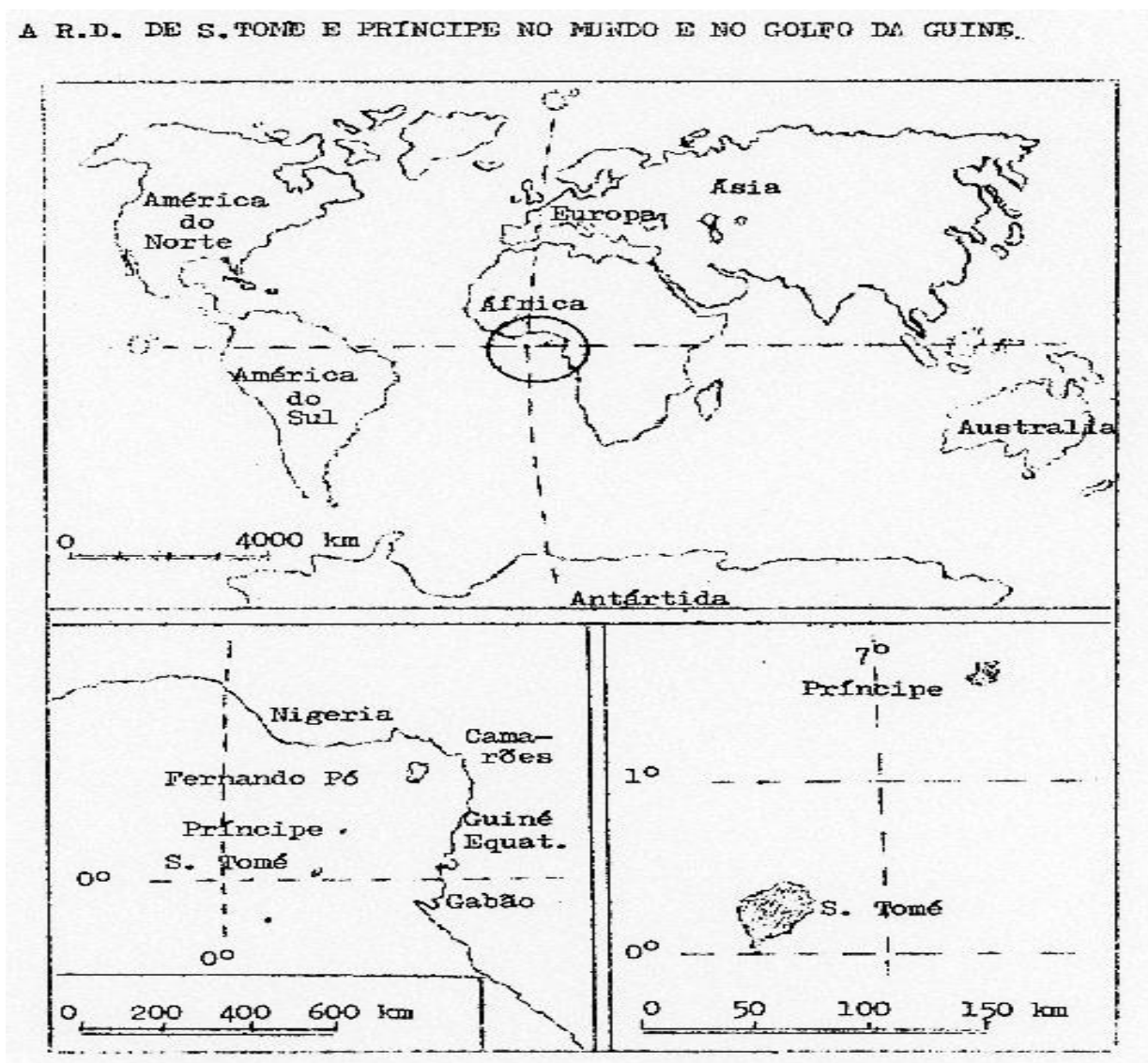
CAPÍTULO 1: SÃO TOMÉ E PRÍNCIPE

1.1. CONTEXTUALIZAÇÃO HISTÓRICA

A República Democrática de São Tomé e Príncipe (RDSTP), ou simplesmente São Tomé e Príncipe (STP), é um país insular situado no Golfo da Guiné a cerca de 400 km da costa ocidental do Continente Africano, bem na linha do Equador, (Latitude: 0.3352, Longitude: 6.7308 0° 20' 7" Norte, 6° 43' 51" Este (Figura 2 p. 26). O arquipélago é formado por duas ilhas principais, Ilha de São Tomé e Ilha do Príncipe, e alguns ilhéus (Ilhéu das Cabras e Ilhéu das Rolas em São Tomé, Ilhéu Bombom, no Príncipe, entre outros). As duas ilhas têm uma área total de 1.001 km². São Tomé está ao largo da costa do Gabão e perto do Equador, com uma área aproximada de 859 km². A ilha do Príncipe se estende a mais de 142 km² e está localizada a 150 quilômetros ao norte de São Tomé.

O arquipélago é de origem vulcânica, muito acidentado e montanhoso (o ponto mais alto atinge 1.500 metros). Possui clima tropical húmido com duas estações (Chuva e Gravana): (i) a estação das chuvas dura nove meses, entre setembro e maio; (ii) Gravana é a estação seca, ocorre entre junho e agosto. Bastante influenciado pelo relevo, a temperatura média anual é de cerca de 26°C no planalto, enquanto a umidade relativa do ar chega a 75%. O isolamento de STP contribuiu para a criação de uma biodiversidade única, formada por uma cobertura florestal terrestre, com florestas de neblina "enrolar", mata seca, cerrado e manguezais, e cobre a maior parte do país (mais de 90%) e a maior parte da população está concentrada na ilha de São Tomé e outra parte, na ilha do Príncipe, apesar de existirem ilhéus (RGPH, 2012). A capital de São Tomé e Príncipe é a Cidade São Tomé, Distrito de Água Grande, na Ilha de São Tomé, e onde está o maior número de habitantes.

Figura 2 - São Tomé e Príncipe no mundo e no Golfo da Guiné



Fonte: Souza (2021. p. 40)

Os navegadores portugueses João de Santarém, Pero Escobar e João de Paiva (Figura 3, p. 27) chegaram a STP no final do século XV. Porém, a data em que isso ocorreu não é consensual entre os historiadores, alguns afirmam que esses navegadores chegaram a São Tomé em 21 de dezembro de 1470 e à Ilha do Príncipe, em 17 de janeiro de 1471, enquanto outros defendem que em 21 de dezembro de 1470 chegaram

à Ilha de São Tomé e, 17 de janeiro de 1472, na Ilha do Príncipe. Originalmente, as Ilhas se chamavam São Tomé e Santo Antão (ou Santo Antônio), por causa dos dias em que os portugueses lá chegaram.

Figura 3 - Forte de São Sebastião/Estátuas dos navegadores portugueses João de Santarém, Pêro Escobar e João de Paiva



Fonte: Ramusel Graça (2018)³

O povoamento efetivo de São Tomé começou em 1493, e o processo de colonização se deu em dois grandes momentos. De acordo com Seibert (2015, p. 100), esse processo de povoamento de STP compreendia duas categorias principais, colonos brancos e africanos escravizados, trazidos inicialmente do Reino do Benim, no Delta do Rio Niger, e desde o início do século XVI também dos Reinos do Congo e do Ndongo (Angola). Ainda segundo o autor:

Os escravos do Benim falavam edo, pertencente às línguas Benue-Congo ocidental, ao passo que os do Congo e de Angola falavam línguas banto (Benue-Congo oriental), kikongo e kimbundu, respetivamente. Essa situação linguística reflete-se na gênese dos três crioulos das ilhas. A mestiçagem genética deu

³ Disponível em: <<https://www.dw.com/pt-002/são-sebastião-a-fortaleza-que-counta-a-história-de-são-tomé/g-46225031>>. Acesso em: 15 mai. 2024.

origem a uma terceira categoria, os mestiços. A alforria de escravos resultou na emergência de uma quarta categoria de negros livres, os chamados forros. Ao longo do tempo, outros escravos alforriados foram sucessivamente integrados na categoria dos forros, os crioulos majoritários de São Tomé. (Seibert, 2015, p. 100)

Essa versão é reforçada por Hagemeijer (2009) no trabalho intitulado “As Línguas de S. Tomé e Príncipe”. O pesquisador argumenta que a história de STP se assemelha à de outros espaços crioulos, como Caraíbas e admite a existência de dois momentos distintos na ocupação de São Tomé: a fase de habitação, que vai do povoamento definitivo em 1493 até aos primórdios da introdução de cana-de-açúcar, por volta de 1520, e a fase de plantação, o período entre 1520 e o fim do século XVI quando o ciclo do açúcar entra em ruptura.

No primeiro momento, houve o povoamento, logrou-se grande prosperidade econômica e o processo de crioulização, nos séculos XV e XVI, que foram seguidos de um declínio econômico e quase abandono pelos portugueses. O segundo momento está vinculado ao colonialismo moderno nos séculos XIX e XX, nos quais o arquipélago recuperou importância estratégica e econômica. Nesse segundo momento, houve restabelecimento da economia de plantação nas ilhas no Golfo da Guiné que trouxe mudanças demográficas importantes consideráveis devido à imigração massiva de trabalhadores contratados de Angola, Cabo Verde e Moçambique, que, durante muitas décadas, ultrapassaram a população nativa em número.

Em outras palavras, a colonização de São Tomé e Príncipe foi caracterizada pelo sistema de *plantation*. A primeira fase foi baseada na monocultura da cana-de-açúcar e utilização de mão-de-obra escrava. A segunda, por sua vez, teve como característica a

monocultura do café no século XVIII e introdução do cacau na segunda metade do século XIX e trabalhadores vindos da África portuguesa, com exceção da Guiné-Bissau.

Para Seibert (2015, p. 101), o objetivo da colonização das ilhas incluía, entre outras coisas, “o estabelecimento de uma colônia de povoamento europeu, a produção de açúcar, a instalação de um entreposto para a navegação marítima para a Ásia e a difusão do cristianismo na região”. Em 1502, o Rei D. João II de Portugal batizou a Ilha de Santo Antônio de “Ilha do Príncipe” em homenagem a seu único filho e herdeiro Afonso, Príncipe de Portugal (1475-1491). A Ilha se transformou numa donataria, com vistas ao povoamento, e ali foi introduzida a cultura da cana-de-açúcar; e somente em 1573, a donataria reverteu à posse da Coroa Portuguesa.

A história de São Tomé e Príncipe está vinculada à história mundial e, direta e profundamente, relacionada à história do Brasil em diversos aspectos. Por isso, é relativamente comum entre os são-tomenses letrados afirmar que a “descoberta” do Brasil constituiu uma desgraça para São Tomé e Príncipe, pois, assim como São Tomé, o Brasil é um país tropical, com condições climáticas semelhantes, mas com dimensões territoriais continentais. Nesse contexto, Seibert afirma:

No período de 1586 a 1613, São Tomé teve 18 governadores, incluídos aqueles eleitos pelo Senado da Câmara (Cunha, 2001). A guerra de mato contra os mocambos também contribuiu para a instabilidade. Durante um dos conflitos entre bispo e governador, em 1595, numa altura em que a indústria de açúcar já entrara em declínio, São Tomé foi abalado por uma grande revolta de escravos liderada por Amador. Durante a revolta, que durou três semanas e em que participaram cinco mil escravos, mais da metade dos cerca de 120 engenhos existentes na altura foram destruídos. Desde os fins do século XVI, as duas ilhas sofreram vários ataques dos franceses e holandeses. De 1641 a 1648, a holandesa Companhia das Índias Ocidentais ocupou a fortaleza e o porto de São Tomé. Todos esses fatores contribuíram para a queda do açúcar em São Tomé. Contudo, o fator principal do declínio do açúcar foi a emergência do Brasil como produtor desse alimento. Os engenhos brasileiros eram mais produtivos, e a qualidade do açúcar, muito superior à de São Tomé. Devido à humidade, o açúcar de São Tomé era pouco consistente e escuro, enquanto o brasileiro era firme, bem refinado e branco. Contudo, o açúcar são-tomense era barato, o que

garantiu a sua venda enquanto o Brasil não produziu em grande escala. Além disso, no Brasil o clima era menos insalubre para os europeus e houve menos instabilidade política do que nas ilhas no Golfo da Guiné. Consequentemente, muitos fazendeiros se mudaram para o Brasil, onde reproduziram o sistema de *plantation*. (Seibert, 2015, p. 107)

1.1.1. Os Angolares

Com relação a todos os temas relacionados à chegada dos portugueses, o povoamento, o sistema de *plantation*, conflitos e as lutas até a independência de São Tomé e Príncipe, nada é mais controverso e inconclusivo do que a origem dos Angolares, que são falantes da língua homônima – Angolar (*Ngola*). Existem três versões para a história: i) Angolares são sobreviventes do naufrágio de um navio negreiro que teria partido de Angola e naufragado perto de São Tomé e que conseguiram nadar até a ilha; ii) Angolares são descendentes de escravizados que fugiram do trabalho forçado e se refugiaram no ôbô⁴, formando quilombos, mais concretamente na região conhecida hoje como São João dos Angolares; e iii) os Angolares são povos autóctones e já habitavam a Ilha de São Tomé antes da chegada dos portugueses.

Para Seibert (2004, p. 47),

O último capitão dos angolares, Simão Andreza, morreu no início do século XX sem deixar descendentes. Da ocupação do seu território resultou a parcial dissolução da organização social, uma maior dispersão da população pela ilha e um processo de aculturação aos padrões culturais dominante dos forros. Apesar da assimilação com os outros crioulos, os angolares conseguiram preservar a sua própria língua e outras particularidades, como as suas habitações sem estacas e as suas aldeias compactas, permanecendo um grupo distinto com a sua própria identidade sociocultural. Atualmente são cerca de dez mil angolares, vivendo nas zonas do litoral sul desde a Ribeira Afonso até Porto Alegre no distrito de Caué e no litoral noroeste desde Neves até Bindá no distrito de Lembá. Além disso existem pequenos grupos perto da cidade de São Tomé em São João da Vargem, Pantufo e Praia Melão. Todas as suas aldeias têm um chefe que se considera descendente dos fundadores da povoação. (Seibert, 2004, p. 47)

⁴ Ôbô é termo emprestado do *Santome* e significa floresta virgem, muito densa e quase impenetrável.

A fuga de escravos sempre ocorreu desde a fase inicial da colonização, mas Seibert (2015, p. 106) afirma que o aumento da população escrava e o endurecimento do regime de trabalho nas *plantations* impulsionaram essa evasão. Além disso, houve fuga de 670 (5%) de 12.904 escravos importados pela Fazenda Real entre 1514 e 1527. O pesquisador argumenta que o relevo montanhoso da ilha vulcânica e a uma densa floresta tropical no interior de São Tomé proporcionaram as condições favoráveis para estabelecimento de pequenas comunidades no sul de São Tomé, fora do controle das autoridades coloniais no Norte. Há relatos de assaltos frequentes a fazendas para roubar alimentos e escravas em razão do déficit de mulheres entre os Mocambos, grupos de escravos autolibertados. E isso fez com que as autoridades criassem uma milícia por volta de 1530, majoritariamente de escravos, que travou conflitos armados contra os mocambos por décadas. O pesquisador cita documentos do século XVIII, em que os descendentes dos escravos fugitivos são chamados “angolis” ou “angolas”, em referência à suposta origem e que, a partir do início do século XIX, ficaram conhecidos por “angolares”.

Para entendermos as diversas narrativas sobre a origem dos Angolares é importante compreender vários contextos históricos que as envolvem e o Rei Amador. Geralmente, os relatos do colonizador – Portugal – parecem favorecer a versão de naufrágio por três motivos: *i*) reconhecer a fuga constante dos escravizados poderia significar a falta de controle e poder sobre eles perante a Coroa); *ii*) reconhecê-los como autóctones faria cair por terra a ideia de “descobridores” e, portanto, donos da(s) ilha(s), tornando-os conhecidos como “invasores”; e *iii*) usar como argumento a semelhança entre os nomes Angola e Angolares. Além disso, questionam até mesmo, por vezes, a

existência de um autoproclamado rei, Rei Amador, Amador Vieira, Rei dos Angolares, Rei dos São-tomenses (ver Figura 4, p. 32).

Figura 4 – Estátua de Rei Amador na Cidade de São Tomé



Fonte: Lourenço da Silva (2018)⁵

Para Carlos Neves, historiador e político são-tomense, a tese de naufrágio não se sustenta e defende os Angolares como autóctones das Ilhas de São Tomé e Príncipe. Em 1975, “Esboço Histórico das Ilhas de S. Tomé e Príncipe”, obra anônima atribuída a ele, Carlos Neves rejeitou a tese do naufrágio afirmando a presença dos angolares antes da chegada dos portugueses. Citando Neves, Seibert afirma:

⁵ Disponível em: <https://www.stp-press.st/2018/11/02/presidente-inaugura-estatua-do-rei-amador-figura-embematica-da-historia-sao-tomense/>. Acesso em: 15 mai. 2024.

Na minha opinião, os Angolares são uma ramificação dos Bantos, que provavelmente se teriam fixado nas regiões do Gabo e do Rio Muni e que posteriormente se tivessem deslocado para algumas das Ilhas do Golfo da Guiné” (Esboço 1975: 23). Também aponta que os próprios Angolares nas suas lendas não se referem a nenhum naufrágio dos seus antepassados. Pergunta-se: “Admitindo que (os Bantos) tenham navegado até Fernão Pó, porque não teriam avançado um pouco mais, até S.Tomé?” Este autor é o primeiro que levanta a questão de como os náufragos podiam vencer a distância entre o rochedo de Sete Pedras e a costa, concluindo que “teriam que ser indivíduos habituados ao mar ou a grandes rios, pois para se salvarem, era indispensável que soubessem nadar...” (Esboço 1975: 22). Dez anos mais tarde um manuscrito com o título “A História da República Democrática de São Tomé e Príncipe”, elaborado em conjunto por um grupo de são-tomenses proeminentes e uma equipa russa, que não chegou a ser publicado, confirma a teoria apresentado no Esboço Histórico. Este manuscrito diz: “O território e a população pequenos das ilhas fizeram com que os autóctones fossem liquidados, dispersos ou expulsos pelos invasores para as regiões não exploradas, no primeiro período, pelos europeus. Os colonialistas queriam apagar a memória dos primeiros habitantes das ilhas e declararam-nas desertas. Mas em breve os portugueses teriam de convencer-se da presença das pessoas que apareceram em S.Tomé independentemente deles e sentir por experiência própria a sua cólera e a intransigência para com a subjugação. Então estas pessoas foram declaradas descendentes dos escravos que se encontravam no navio naufragado junto às costas de S.Tomé nos anos 40 do século XVI, e receberam o nome de “angolares” (História da RDSTP 1985: 274). Nos anos oitenta o jornalista português Jorge Trábulo Marquês foi de canoa de Tomé para o Príncipe e para a Nigéria a fim de provar que os antepassados dos Angolares podiam ter chegado ao sul de S.Tomé através do mar nas suas próprias embarcações. (Seibert, 1998, n.p.)

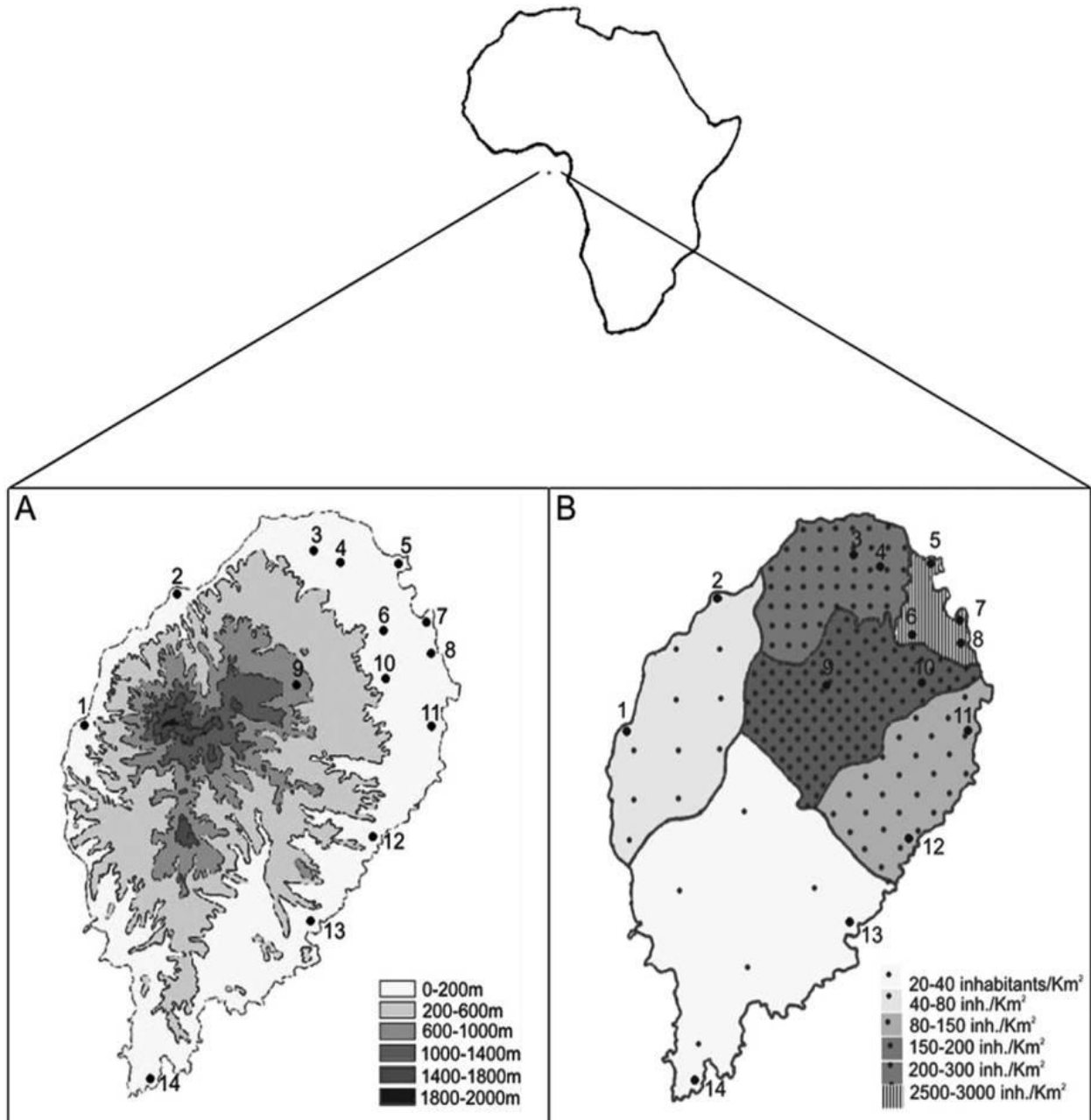
Além disso, Coelho et al. (2008) no artigo *Human Microevolution and the Atlantic Slave Trade: A Case Study from São Tomé*, apresentam informações muito relevantes com os estudos sobre a genética dos grupos populacionais de São Tomé. Segundo os autores, havia populações autóctones quando os navegadores portugueses desembarcaram em São Tomé e Príncipe, pois o DNA dos Angolares é diferente do dos povos escravizados levados a STP pelos colonizadores. De acordo com Coelho et al. (2008, p.135), houve um erro metodológico nos estudos genealógicos das populações da Ilha de São Tomé. Eles argumentam:

Estudos genéticos anteriores em São Tomé concentraram-se principalmente na análise de marcadores uniparentais, utilizando várias abordagens de amostragem. Mateus et al. (1997) caracterizaram a variação da sequência do mtDNA tratando a ilha como uma unidade de amostragem única. Trovada et al. (2001, 2003) estudaram os microssatélites do cromossomo Y e o mtDNA entre doadores autodeclarados de grupos etnolinguísticos preconcebidos. Buscamos aqui descrever e interpretar a estrutura genética de São Tomé sem depender de predefinições étnicas, anatômicas ou categorias geográficas. Nosso principal objetivo foi entender como diferentes fatores evolutivos moldaram a atual variação genética em São Tomé, propondo os seguintes questionamentos: Existe algum nível detectável de estrutura populacional em São Tomé? Quantos agrupamentos genéticos diferentes, se houver, podem explicar a estrutura observada? Como eles se originaram, e como eles se relacionam com os recursos linguísticos e informações históricas sobre o povoamento da ilha? (Coelho et al., 2008, p.135) (tradução nossa)

Os pesquisadores constataram que São Tomé se caracteriza por um nível notável de estruturação genética que é mais bem capturada por dois clusters (agrupamentos) que são mais intimamente relacionados à diferenciação linguística do que à distância geográfica. Esses resultados mostram que métodos apropriados para análise de agrupamento genético em regiões muito distantes podem ser úteis na classificação de indivíduos no nível microgeográfico. Coelho et al. (2008) ressaltam ainda que as nossas observações mostram que as análises baseadas em fronteiras dêmicas geneticamente definidas são uma alternativa útil para os estudos dos projetos baseados no uso de grupos pré-concebidos como pontos de partida para comparações genéticas populacionais. Portanto, concluíram que:

Em suma, os nossos resultados sugerem que a atual estrutura do grupo Angolar foi fortemente moldada por um evento fundador de parentesco estruturado, resultante da fuga de um clã patrilinear de escravos rebeldes que estabeleceram contatos secundários com o resto da ilha, principalmente através de fluxo de genes restritos e mediados por mulheres. Apesar desse fluxo genético, o grupo manteve elevados níveis de diferenciação genética que são mais típicos de um grupo centralizado do que de uma federação livre de refugiados. O impacto de um efeito fundador na estrutura genética dos Angolares é também reforçado pelos dados demográficos disponíveis. Hoje, estima-se que os Angolares totalizem cerca de 10.000 pessoas (SEIBERT, 1998), o que corresponde aproximadamente às populações combinadas de São João dos Angolares (4.082), Ribeira Afonso (4.955) e Santa Catarina (2.199), representando cerca de 8% dos 131.633 habitantes de São Tomé [Figura 5, p. 35]. (Coelho et al., 2008, p.141) (tradução nossa)

Figura 5 - São Tomé: posição geográfica no Golfo da Guiné, topografia (A), demografia (B), e locais das amostras (adaptado de Pinto et al. 2003). 1, Santa Catarina; 2, Neves; 3, Guadalupe; 4, Agostinho Neto; 5, Praia Gamboa; 6, Madalena; 7, Cidade de São Tomé; 8, Pantufo; 9, Monte Café; 10, Trindade; 11, Santana; 12, Ribeira Afonso; 13, São João dos Angolares; 14, Porto Alegre.



Locais 4, 9 e 14 são *plantations* (plantações).

Fonte: Coelho et al. (2008, p. 136)

1.1.2. Símbolos Nacionais de São Tomé e Príncipe

Figura 6 - Bandeira de São Tomé e Príncipe

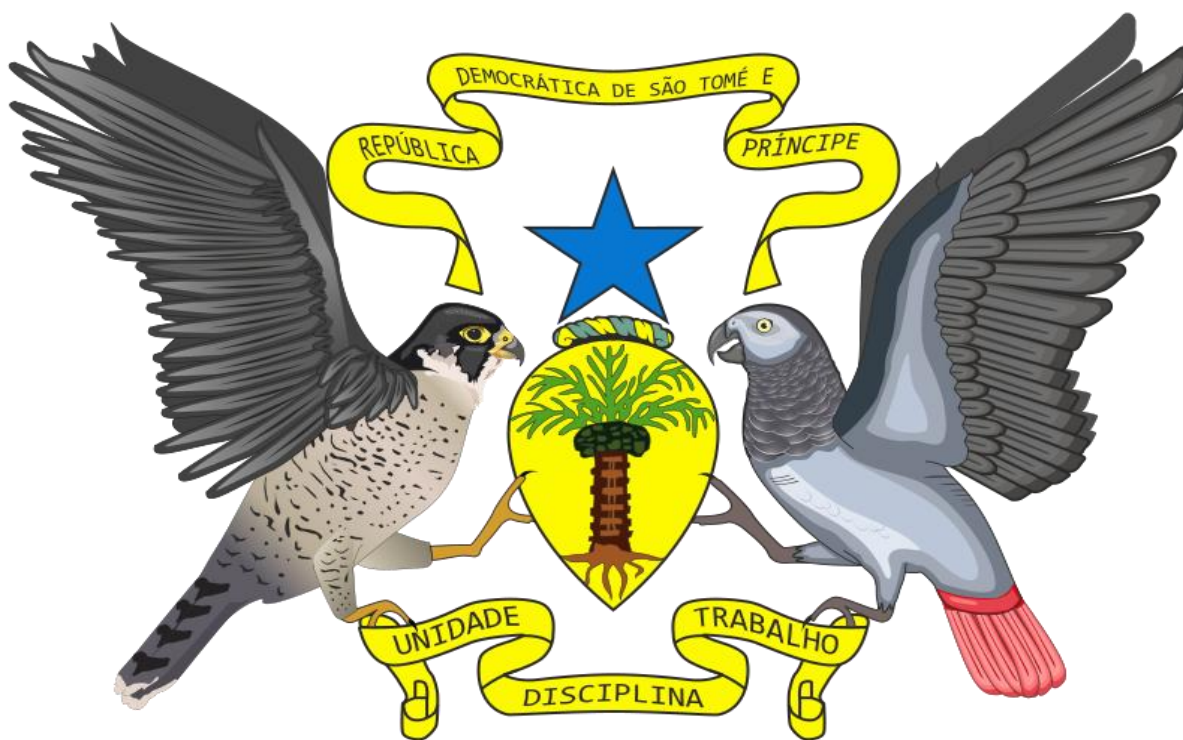


Fonte: Elaborado pelo autor

Após a “independência” em 12 de julho de 1975, São Tomé e Príncipe adotou os seguintes símbolos nacionais: a Bandeira Nacional, o Hino Nacional e a Insígnia (Brasão de Armas). A bandeira de São Tomé e Príncipe (Figura 6, p. 36) foi adotada oficialmente em 5 de novembro de 1975, quatro meses após a independência de Portugal. As cores presentes na bandeira remetem ao Panafricanismo - verde, amarelo e preto. O triângulo vermelho simboliza a longa jornada de luta e sangue derramado até a independência, as duas estrelas negras representam o país duas ilhas principais, em que a cor preta representa o povo negro e ao mesmo tempo o luto por aqueles perderam a vida durante a colonização. O verde, por sua vez, simboliza as florestas e o amarelo, o cacau maduro, que continua sendo um dos melhores do mundo, embora produzido em pequena quantidade.

Outro símbolo nacional - o hino de São Tomé e Príncipe, intitulado “Independência Total” é de autoria da falecida escritora, poetisa, professora e política são-tomense, Alda do Espírito Santo, com música de Manuel dos Santos Barreto de Sousa e Almeida e foi adotado logo após o país conquistar a independência de Portugal, em 12 de julho de 1975. Porém, passados 48 anos da independência, São Tomé e Príncipe parece ainda não ter realizado o tão almejado sonho idealizado pela poetisa e inscrito no título do seu hino nacional – independência total, já que continua dependente de ajuda externa.

Figura 7 - Brasão de Armas de São Tomé e Príncipe



Fonte: Wikipedia (2023)⁶

⁶ BRASÃO DE ARMAS DE SÃO TOMÉ E PRÍNCIPE. In: WIKIPÉDIA, a enciclopédia livre. Flórida: Wikimedia Foundation, 2021. Disponível em: https://pt.wikipedia.org/w/index.php?title=Bras%C3%A3o_de_armas_de_S%C3%A3o_Tom%C3%A9_e_Pr%C3%ADncipe&oldid=62712776. Acesso em: 21 mai. 2024.

A insígnia ou brasão de armas de São Tomé e Príncipe (Figura 7, p. 37), constituída pela figura de um falcão à esquerda e um papagaio à direita, separados por um brasão de forma ovular, cuja abcissa vertical, é de dimensão 0,33 vezes superior que a horizontal e no interior se destaca uma palmeira situada ao longo da abscissa vertical. Acima, existe uma faixa que indica o nome do país. Na base dos braços o lema nacional, “a unidade, a disciplina de trabalho”. (São Tomé e Príncipe, 2003 [1975], p. 5)

Ambas as aves foram escolhidas por serem muito comuns em STP e representativas das duas ilhas, sendo o falcão em São Tomé e o papagaio, no Príncipe. Além disso, o papagaio é tão marcante e característico da ilha do Príncipe que está presente na antroponímia da ilha – Ilha de Papagaio, Rio Papagaio, Pico de Papagaio etc. Vale ressaltar que, antes de conquistar a autonomia e passar a se chamar RAP (Região Autônoma do Príncipe), Príncipe era apenas mais um distrito de São Tomé e Príncipe, o Distrito de Pagué. E “pagué”⁷ [pa'gɛ] nada mais é que papagaio em Lung'le. Além desses símbolos, apresentamos a seguir uma nota de cinquenta mil dobrás (DBs), a moeda de São Tomé e Príncipe (Figura 8, p. 39), cujo código internacional era STD (*Sao Tome and Príncipe Dobra*), foi renomeado em 2017 e passou em 2018 a se chamar STN (*New Sao Tome and Príncipe Dobra*).

⁷ Em Santome, papagaio é *papage* [papa'gɛ].

*A ortografia original *pagué* foi mantida (e para Santome seria *papagué*), pois antecede o Decreto 19/2013, que fixou a escrita para as Línguas Nativas de São Tomé e Príncipe.

Figura 8 - Nota de Dbs. 50.000



Fonte: Banco Central de São Tomé e Príncipe (2004)

Nesta figura destacam-se o herói nacional, Rei Amador, rosto presente em todas as cédulas emitidas naquela altura, assim como o pássaro endêmico das Ilhas de São Tomé e Príncipe, nacionalmente conhecido como *Conóbia*, palavra da língua Santome, cujo nome científico é *Alcedo nais*. E assim era popularmente conhecida e chamada a nota de 50.000Dbs – *Conóbia*⁸.

Embora essa metonímia não seja exclusiva do Português de São Tomé (PST), ressaltamos a importância e o impacto da língua Santome (ST) na formação do léxico e da gramática do português falado em São Tomé e Príncipe (STP). O uso de parte da imagem para se referir ao todo, sinédoque, e o fato dessa ave assim como muitas outras aves e plantas serem, originalmente, nomeadas no ST demonstram a tamanha contribuição do ST, por meio da fauna e da flora de STP, para o PST e formação de sentidos no uso da língua no quotidiano pelas comunidades de fala nos mais variados

⁸ Embora esta tese adote o ALUSTP, que determina que o fonema /k/ seja representado pelo grafema “k” e que as vogais médias abertas não recebam qualquer acentuação, o item lexical *Conóbia* antecede o ALUSTP, sua grafia está consolidada e é oficialmente usada. Portanto, para esse caso, adotaremos a grafia oficial *Conóbia*, e não *Konobyia*.

contextos, como a arte, a música, provérbios, adivinhas e manifestações culturais de STP (ver exemplo 9 e 10).

(9) **Advinhas**

a) *N da son, mundu jinga.*ST

1SG dar chão mundo tremer

Eu caí e o chão tremeu.^{PT}

Resposta: izaquente⁹

b) *N da son, wê lanka.*ST

1SG dar chão olho arrancar

Eu caí e o olho arrancou.^{PT}

Resposta: andim¹⁰

Importa ressaltar também que essas adivinhas (9a e 9b) são tipicamente feitas no ST, mas as respostas são dadas geralmente em PST, e de vez em quando em ST. No entanto, os itens lexicais das respostas – izaquente^{PST} e andim^{PST} – têm origem no ST

⁹ *Treculia africana*, árvore tropical, da família das Moráceas, existente em São Tomé, tem raiz medicinal e produz frutos esféricos, grandes e pesados, de cor amarelada, com abundantes sementes comestíveis que se utilizam no fabrico de uma farinha alimentar (Porto Editora). Disponível em: <https://www.infopedia.pt/dicionarios/lingua-portuguesa-aa/izaquente>. Acesso em: 22 mar. 2024.

¹⁰ *Elaeis guineensis* palmeira de origem africana, de espique simples, folhas grandes, penatissetas, dispostas em coroa terminal, flores monoicas, reunidas em espádices interfoliáceos e fruto drupáceo, de formato ovoide e cor vermelha, alaranjada ou amarela, de que se extrai óleo de duas qualidades (da polpa ou da semente), usados na alimentação, na indústria, etc. (Porto Editora). Disponível em: <https://www.infopedia.pt/dicionarios/lingua-portuguesa/andim>, acesso em 22 mar. 2024.

zêkêntxi / jêkêntxi e *andji*. Nesse sentido, estamos diante tanto de *code-switching* como de lexicalização.

(10) Provérbios

a) *Pema pô sa nglanji. Xi kodo buya e, ndala ka dêsê.*ST

Palmeira poder ser grande. Se corda abraçar 3SG folha ASP descer

Palmeira poder ser grande. Se a corda a abraçar, a andala desce.^{PT} (*lit.*)

A palmeira pode ser grande. Se a corda der volta nela, o vianteiro¹¹ sobe.^{PT}

b) *Pema môlê, Tonha be dê.*ST

Palmeira morrer Antônio ir dela

Fogo acabou, a mulher abandonou.^{PT}

Em ambos os casos, observamos o uso do item lexical *pema* [palmeira], mas no exemplo 10b, a palavra é usada de forma metafórica – o significado de uma palavra de seu contexto convencional (denotativo) foi transferido para um novo campo de significação (conotativo). Em 10b, *pema* [palmeira] é usada com sentido de órgão genital masculino. Ressaltamos que esses provérbios são usados exclusivamente em ST, mesmo que a interação esteja sendo feita em PST. Portanto, verificamos ocorrência de

¹¹ Vianteiro^{PST} (do Santome *Vyantêlu* [vjẽ'telu]) (Figura 9) é um profissional de sexo masculino que sobe e limpa a palmeira, colhe o andim^{PST} (dendê^{PB}) e extrai o vinho da palmeira chamado de “Vinho da Palma”. Além do Vianteiro, existe também o Cortador de Andim^{PST} (emprestado também do Santome *Kotadô d'anji*), cuja função é basicamente a mesma, exceto a extração do vinho da palma, pois esta última exige habilidades específicas. Embora exista o item lexical “vinhateiro” no português, o comum em STP é Vianteiro.

code-switching intersentencial tanto nos casos dos exemplos 9a e 9b quanto em 10a e 10b.

Figura 9 – Vianteiro



Fonte: Viagem a São Tomé (2023)¹²

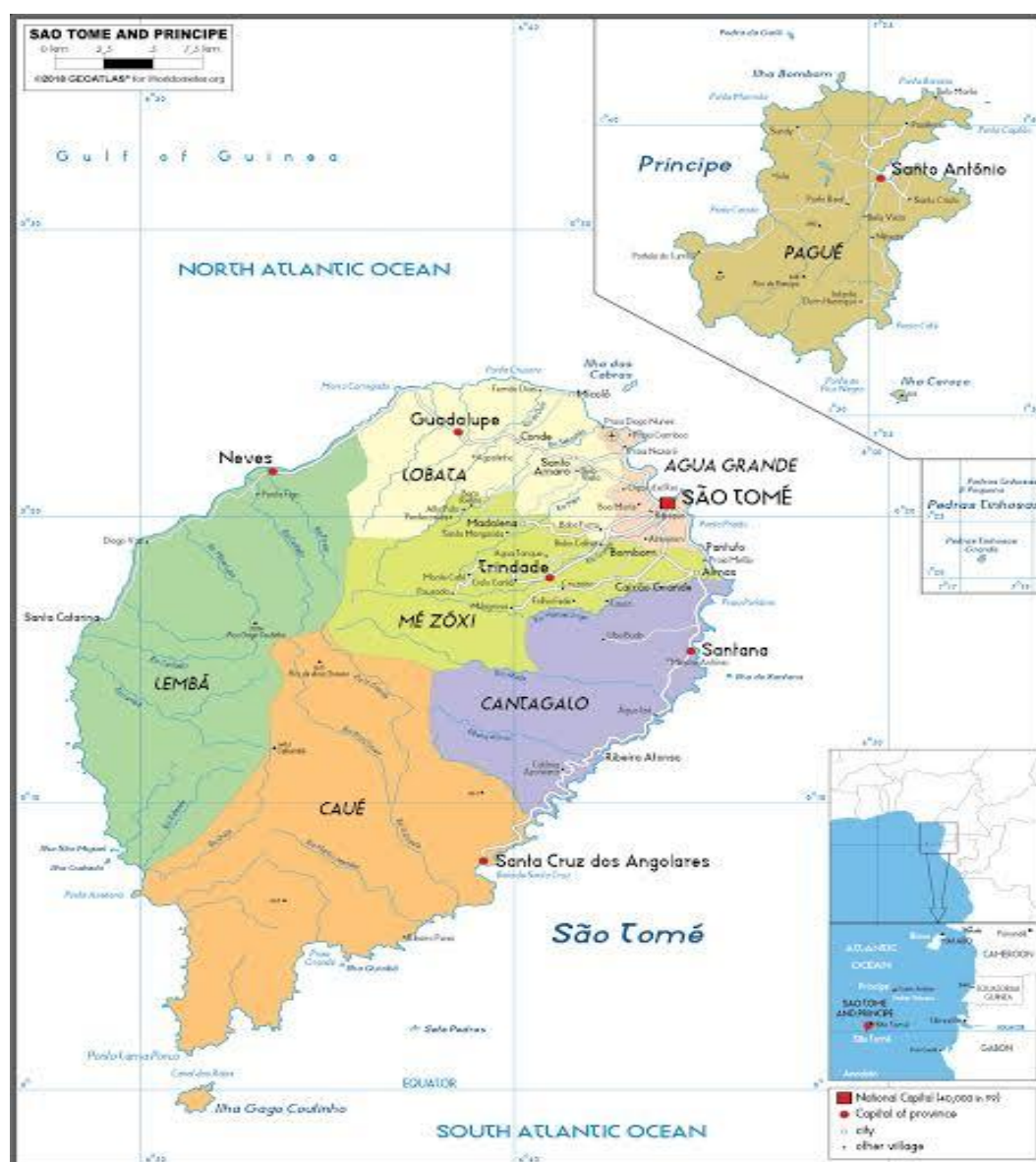
¹² Vianteiro. Disponível em: <https://www.facebook.com/viagemasaotome/posts/extraindo-o-fabuloso-delicioso-vinho-da-palmasaotome-beach-tours-viagemeturismo-/736450475171742/>. Acesso em: 21 set. 2025.

1.2. CONTEXTO ATUAL

Hoje, com 48 anos de independência “protocolar”, já que ainda depende econômica e financeiramente de Portugal e outros parceiros internacionais, São Tomé e Príncipe é um país multipartidário. Foram quinze anos de monopartidarismo do Movimento de Libertação de São Tomé e Príncipe/Partido Social-Democrata (MLSTP/PSD) até a abertura política no começo da década de noventa.

Do ponto de vista administrativo, o país é dividido 7 distritos (Figura 10, p. 44), agrupados em 4 Regiões: a Região Norte (Distritos de Lobata e Lembá), a Região Centro (Distritos de Água Grande (Capital) e Mé-Zóchi), a Região Sul (Distritos de Caué e Cantagalo), e a Região Autónoma do Príncipe (RAP). Os distritos, por sua vez, estão subdivididos em cidades, vilas e localidades. Em termos de superfície, o distrito de Caué é o maior, ocupando uma superfície de 267 Km², seguido de Lembá com uma superfície de 229 Km², Mé-Zochi com 122 Km², Cantagalo com 119 Km², Lobata com 105 Km² e Água Grande com uma superfície de 16,5 Km². Os distritos são representados por uma autarquia (Câmara Distrital), dirigida por um presidente eleito democraticamente. Diferente dos demais distritos, o distrito de Pagué conquistou, em 1996, o status de Região Autónoma do Príncipe e passou a ser dirigida por um Governo Regional chefiado por um presidente, eleito democraticamente (INE-STP, 2014).

Figura 10 - Mapa político de STP



Fonte: Worldometers (2023)¹³

Dados do Recenseamento Geral da População e Habitação (RGPH), feito pelo Instituto Nacional de Estatística de São Tomé e Príncipe (INE-STP) em 2012, publicado em 2014, mostram que a população estimada para 2017 seria de 197.700 habitantes,

¹³ Disponível em: <https://www.worldometers.info/maps/sao-tome-and-principe-political-map/>. Acesso em: 22 mai. 2024.

sendo 49% homens e 51% mulheres. A taxa de crescimento anual é de aproximadamente 2%. Desta população, 60% vivem num raio de 10 km do centro da capital do País, que é a cidade de São Tomé (Figura 11, p. 45). Os conglomerados populacionais estão fundamentalmente concentrados em pequenas cidades, vilas e localidades, empresas agropecuárias e aldeias pesqueiras. Os distritos mais populosos são Água Grande e Mé-zóchi, representando 64% da população total do país. Porém, segundo dados do Banco Mundial, em 2019, STP já tinha uma população de 215.000 habitantes.

Figura 11 - Baía de Ana Chaves / Cidade de São Tomé



Fonte: Wikipedia (2023)¹⁴

Ainda segundo dados do INE-STP (2014, p. 9):

A sua [da população] repartição por meio de residência mostra que 67% residem no meio urbano e 33% no meio rural. Esta tendência é verificada também a nível distrital, com excepção dos distritos de Mé-Zochi, de Lobata e da Região

¹⁴ SÃO TOMÉ (SÃO TOMÉ E PRÍNCIPE). In: WIKIPÉDIA, a enciclopédia livre. Flórida: Wikimedia Foundation, 2023. Disponível em: <[https://pt.wikipedia.org/w/index.php?title=S%C3%A3o_Tom%C3%A9_\(S%C3%A3o_Tom%C3%A9_e_Pr%C3%ADncipe\)&oldid=67140805](https://pt.wikipedia.org/w/index.php?title=S%C3%A3o_Tom%C3%A9_(S%C3%A3o_Tom%C3%A9_e_Pr%C3%ADncipe)&oldid=67140805)>. Acesso em: 19 dez. 2023.

Autónoma do Príncipe onde existe claramente maior concentração da população no meio rural. Entre os dois recenseamentos, a densidade populacional à nível do país teve um aumento na ordem dos 31,8%, passando de 135,5 habitantes por quilómetros quadrados em 2001 para 178,6 habitantes por quilómetros quadrados em 2012. O distrito de Água Grande é o que sofre maior pressão humana sobre o seu território (4209,3 hab/Km²), ao contrário do distrito de Caué que possui a mais baixa concentração da população (22,6 hab/Km²). A estrutura da população Santomense é caracterizada pela predominância da população jovem e relativamente escassa população idosa. Com efeito, a população residente com menos de 20 anos corresponde a 52,1% da população total e 3,8% têm mais de 65 anos de idade. (INE-STP, 2014, p. 9)

As características geográficas e populacionais de São Tomé e Príncipe são muito importantes para os estudos linguísticos locais e não só. Portanto, a inclusão e análise desses dados populacionais são muito importantes para a presente pesquisa porque essas migrações internas e evolução da população apresentadas nos Quadros 1 e 2 (pp. 18 e 20, respectivamente) e Quadro 3 (p. 46), a seguir, impactam diretamente e de forma determinante o contexto, as dinâmicas sociais e as línguas faladas no arquipélago, ativando ou acelerando, portanto, os processos de mudança linguística.

Quadro 3 - População residente segundo o distrito de residência atual por local de nascimento

Local de Nascimento	Distrito de residência atual							TOTAL
	Água Grande	Mé-Zóchi	Cantagalo	Caué	Lembá	Lobata	Príncipe	
Água Grande	33.126	2.775	570	262	503	1.067	530	38.833
Mé-Zóchi	7.490	27.764	876	223	568	963	339	38.223
Cantagalo	2.934	1.526	10.315	224	280	405	173	15.857
Caué	1.067	528	470	4.370	419	156	70	7.080
Lembá	1.728	941	518	271	8.368	854	353	13.033
Lobata	2.971	805	240	76	332	11.324	128	15.876
Príncipe	1.121	269	77	18	60	140	4.320	6.005
Exterior do país	1.449	497	192	57	166	278	53	2.692
TOTAL	51.886	35.105	13.258	5.501	10.696	15.187	5.966	137.599

Fonte: INE-STP (2003b, p. 24)

De acordo com a Constituição da República Democrática de São Tomé e Príncipe de 10 de setembro de 1990, revista em 2003 e em vigor desde 2006, o país é uma democracia parlamentar com um sistema semipresidencialista em que o poder executivo é exercido pelo primeiro-ministro, o chefe de Governo, e compartilha com o Presidente da República certas esferas de poder, como relações internacionais, diplomacia e defesa nacional.

Segundo dados do Governo São-tomense,

A economia são-tomense, como a de outros Pequenos Estados Insulares em Desenvolvimento (PEID), é altamente limitada pela insularidade do país, sua fragilidade e pelos recursos limitados. Isso resulta em extrema vulnerabilidade a choques externos e forte dependência da Ajuda Pública ao Desenvolvimento (APD), que financia mais de 90% das despesas de investimento (média de 93,6% em relação ao período de 2012-2015). O sector terciário, em grande parte informal, representa quase 60% do PIB e emprega 60% da população ativa, enquanto os sectores primário e secundário contribuem cada um com 20% do PIB, de acordo com dados mais recentes. Durante o período de 2010 a 2014, o país registou uma tendência positiva no crescimento económico, com uma taxa média de crescimento de 4% do PIB. De acordo com as projeções do FMI, este crescimento irá oscilar entre 5% e 9% entre 2015 e 2020. Segundo o Banco Mundial, São Tomé e Príncipe tem um Produto Interno Bruto (PIB) per capita de 1730 USD (estimado em 2016). São Tomé e Príncipe tem um IDH para 2017 de 0,589 - o que coloca o país na categoria de desenvolvimento humano médio - posicionando-o em 143 de 189 países e territórios. Apesar do desempenho macroeconómico dos últimos anos, os níveis de pobreza e a desigualdade social mantêm-se como problemas a serem ultrapassados. De acordo com o Inquérito ao orçamento familiar (IOF) realizado em 2010, cerca de 66,2% da população vive numa situação de pobreza absoluta, afetando mais as mulheres (71.3%) do que os homens (63.4%). Os distritos com taxas de pobreza superiores a 70% (Caué, Lembá, Príncipe) só explicam 19% da pobreza nacional. Em contrapartida, os distritos de Água Grande, de Mé-Zóchi e de Lobata que albergam 74,1% da população total do país, contribuem com 71,2% da pobreza nacional. O aumento da população e o estado de pobreza levaram as populações rurais a migrarem para as zonas urbanas. (Ministério da Saúde-STP, 2018, p. 14)

1.3. LÍNGUAS

1.3.1. Línguas Nativas de São Tomé e Príncipe

Em vista da história, povoamento e processo socioeconômico e político, São Tomé e Príncipe tem atualmente três línguas nativas - Santome (ou Forro) e Ngola¹⁵, em São Tomé, e Lung'le, na Ilha (de Santo Antônio) do Príncipe - além do Kabuverdianu e Tonga, e o Português, a única língua oficial. Assim sendo, o Português era a única língua formalmente ensinada nas escolas como língua materna (L1). Apesar de algumas tentativas não bem-sucedidas da formalização do ensino do Lung'le na ilha do Príncipe e a intenção política de fazer o mesmo com o Santome em São Tomé, o projeto finalmente se concretizou. Em 2021, com o apoio da cooperação brasileira através da Embaixada do Brasil em São Tomé e dos pesquisadores brasileiros, Ana Livia Agostinho (UFSC) e Gabriel Antunes de Araújo (USP), foi oficialmente lançado o livro *Lung'le, lunge no: método para aprender lung'le*. E a partir do mês de novembro de 2023, Lung'le está oficialmente incluído no currículo escolar da Ilha do Príncipe com material didático-metodológico adequado para o ensino-aprendizagem, com o objetivo de preservar e promover a língua, cultura e a identidade regional e nacional.

Embora o povoamento efetivo de São Tomé e Príncipe tenha iniciado em 1493, a primeira referência histórica que sugere a existência efetiva de uma língua crioula no Golfo da Guiné é bem posterior, datando de 1627, época em que o Padre Sandoval

¹⁵ Quando um Forro usa frases como “Fulano é ngola (Angolar), tonga (Tonga), moncó (Principense) ou nhõ (Kabuverdianu)”, faz referência às pessoas, não à língua, e geralmente com conotação pejorativa. E o propósito mencionar o uso de “Moncó” não é criar constrangimento ao Forro, mas pontuar a questão e esclarecer o item lexical “moncó”, uma vez que se verificou a sua utilização de forma inapropriada em alguns trabalhos de pesquisa linguística sobre São Tomé e Príncipe.

escreve em Cartagena (Colômbia). Citando Sandoval, Germán de Granda (1970, p. 2)

afirma:

Y los que llamamos criollos y naturales de San Thomé, con la comunicación que con tan bárbaras naciones han tenido el tiempo que han residido en San Thomé, las entienden casi todas con un género de lenguaje muy corrupto y revesado de la portuguesa que llaman lengua de San Thomé, al modo que ahora nosotros entendemos y hablamos con todo género de negros y naciones con nuestra lengua española corrupta, como comúnmente la hablan todos los negros. (Granda, 1970, p. 2)

Em *A Corpus of Santome*, Hagemeyer et al. (2012, p. 61) afirmam que:

Santome (São-Tomense, Forro) é um crioulo de base portuguesa falado na ilha de São Tomé no Golfo da Guiné (África). A língua não tem status oficial, mas segundo o censo de 2001, 72,4% da população com mais de 5 anos de idade falava Santome (como L1 ou L2), num total de 137.599 habitantes (RGPH – 2001, 2003). O Português, a língua oficial e mais falada na ilha, a porcentagem era de 98,9%, o que mostra que existe um alto grau de bilinguismo. No entanto, os são-tomenses, particularmente as gerações mais novas, têm usado português cada vez mais. O resultado é uma perda gradual da diglossia e atrito linguístico do crioulo. (Hagemeyer et al., 2012, p. 61) (tradução nossa)

Hagemeyer (2009, p. 4) destaca que as mulheres africanas e os filhos que nasciam dos casamentos e concubinatos com europeus foram oficialmente declarados livres a partir de 1515 e 1517, respectivamente, e rapidamente constituíram uma comunidade com reivindicações e poderes socioeconômicos próprios. Assim sendo, o autor reforça a plausibilidade de que esta comunidade de forros, escravos que recebiam a carta de alforria, com uma identidade própria, tenha estado na origem e consolidação da nova língua que se falava na ilha. Dessa forma, o crioulo ter-se-á rapidamente difundido para as roças, no regime de plantação, tornando-se a língua-alvo dos escravos recém-chegados para efeitos de comunicação.

A “carta de alforria” é, segundo Karasch (2000, p. 439), a “prova da liberdade de um escravo, introduzindo-o na vida precária de uma pessoa liberta numa sociedade escravista”. Hagemeyer (2009, p. 4), por sua vez, conceitua “forro” como “escravos que

recebiam a carta de alforria, afirmando que “a palavra ‘forro’ (*fôlô* em crioulo), o nome alternativo para o Santome, é derivada de ‘(carta de) alforria’”. No entanto, para um são-tomense (nativo do país São Tomé e Príncipe), “forro” geralmente não remete à noção de escravo ou escravizado. Pelo contrário, a noção de “forro” além da língua nativa oficialmente designada de Santome, é de indivíduo natural da Ilha de São Tomé, dono da terra, tido como o são-tomense mais autêntico e originário da Ilha, que não se submete às ordens dos colonizadores ou de outrem. Para reforçar esse conceito de são-tomensidade, a expressão *fôlô jikiti*, no Santome, ou forro jikiti, no PST, pode ser traduzida como são-tomense puro ou forro puro.

De acordo com Gonçalves & Hagemajjer (2015, p. 87), o Santome deriva diretamente do protocrioulo de base portuguesa, resultado do pidgin que se formou do contato entre os portugueses e africanos que já habitavam a Ilha de São Tomé no século XVI. Ainda segundo os autores, o Ngola, que é falado, principalmente, na Cidade de São João dos Angolares e o Lung’lé derivaram do Santome:

Do contacto entre os portugueses e os africanos que povoaram a ilha terá surgido um *pidgin* capaz de assegurar a comunicação básica num contexto multilíngue. A rápida expansão e nativização deste *pidgin* deu origem a um crioulo de base lexical portuguesa, que passou a ser a língua da comunidade de escravos (livres). A continuação no tempo desta proto-língua é a língua que hoje é conhecida como forro, o crioulo maioritário de São Tomé e Príncipe. Em consequência da difusão desta língua para as ilhas do Príncipe e de Annobón e para o interior da ilha de São Tomé no decurso do século XVI, falam-se hoje mais três crioulos para além do forro: o 'lung'ie', o 'fa d'ambô' [na Ilha de Ano-Bom] e o 'angolar' [*Ngola*]. Apesar de descenderem da mesma proto-língua, a inteligibilidade mútua entre estas quatro línguas é hoje limitada. (Gonçalves & Hagemajjer, 2015, p. 87)

O Kabuverdianu, por sua vez, é uma língua que foi transplantada das Ilhas de Cabo Verde para as Ilhas de São Tomé e Príncipe no século XX, já no segundo período de colonização portuguesa em STP, que se inicia com a nova tentativa de

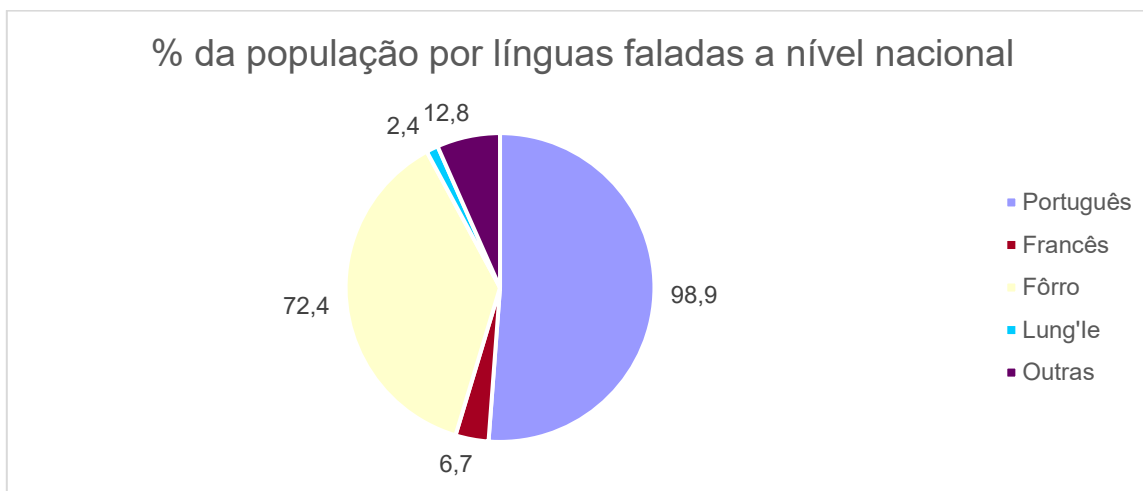
implementação do sistema *plantation*¹⁶, pautado desta vez na cultura do café e do cacau através do regime de contrato de trabalhadores, também chamados de serviçais (Seibert, 2015, p. 110). Por fim, o Tonga é uma língua crioula derivada do contato do português e as línguas africanas, geralmente falado pela comunidade descendente dos trabalhadores contratados (os chamados “serviçais”), oriundos dos países africanos, principalmente Angola, Cabo Verde e Moçambique.

Dados do censo demográfico de São Tomé e Príncipe de 2001 (RGPH-2001), publicados em 2003, indicam o Português, Forro [Santomé], Lunguié e o Francês como as principais línguas faladas no arquipélago. De acordo com o INE-STP (2003a, p. 53), 98,9% da população falava o português com distribuição uniforme entre as faixas etárias e sexo em todo o território nacional, tanto urbano quanto rural. O Santome (ou Forro) era a segunda língua mais falada com 72,4% e o Lung’le com apenas 2,3% de falantes, e francês com 6,7%, sendo a língua estrangeira mais falada em razão da localização geográfica de STP em relação aos países francófonos do Golfo da Guiné, principalmente o Gabão (ver Gráfico 1, p. 52).

Ainda de acordo com o censo de 2001, “além do Príncipe, o distrito onde se fala mais o Lunguié é Água Grande, por ser onde se encontra a capital, e provavelmente, onde se podem encontrar muitas pessoas oriundas da ilha do Príncipe”. Uma curiosidade apresentada pelo censo é o fato do Ngola ser falado por 47,8% da população da RAP.

¹⁶ Segundo Seibert (2015, p. 105), *plantation* era “um tipo de produção agrícola em larga escala, baseada em trabalho escravo, monocultura e exportação”.

Gráfico 1: Proporção (%) da população por línguas faladas a nível geral



Fonte: INE-STP (2003a, p. 53)

1.3.2. Alfabeto Unificado das Línguas Nativas de São Tomé e Príncipe (ALUSTP)

Para preservar e promover as línguas nativas de São Tomé e Príncipe, o Governo São-Tomense, através do Decreto Nº19/2013¹⁷ de 27 de maio de 2013, fixou uma escrita para as Línguas Nativas de S. Tome e Principe - *Santome*, *Ngola* e o *Lung'le* e criou o Alfabeto Unificado para as Línguas Nativas de São Tomé e Príncipe (ALUSTP). O ALUSTP é um conjunto de sinais gráficos, de base latina, para a representação uniforme de cada som das Línguas Nativas de São Tomé e Príncipe e é constituído dos seguintes grafemas: minúsculas (a-b-d-db-dj-e-f-g-gb-i-j-k-kp-l-lh-m-n-nh-p-r-s+th-tx-u-v-w-x-y-z) e maiúsculas (A-B-D-DB-DJ-E-F-G-GB-I-J-K-KP-L-LH-M-N-NH-P-R-S+TH-TX-U-V-W-X-

¹⁷ Pode-se notar que o decreto que instituiu o ALUSTP traz, de forma bastante explícita (embora possa não ser nem implícita para leitores ou pesquisadores pouco atentos), uma questão identitária e de posicionamento político em relação ao português - os nomes das três línguas nativas estão oficialmente escritos nas respectivas línguas, não em português.

Y-Z) e admite uso excepcional de outros grafemas para escrita de estrangeirismos utilizados nas línguas nativas são-tomenses. (STP, 2013, p. 661)

Quadro 4 - Vogais orais breves

Fonema	Grafemas	Santome	Lung'le	Ngola
/a/	<a, A>	laba 'lavar'	<i>daman</i> 'amigo'	/ata 'levantar
/ɛ/	<e, E>	mese 'mestre',	<i>upetu</i> 'espeto'	<i>methe</i> 'mestre'
/e/	<ê, Ê>	<i>mêsé</i> 'querer'	<i>upetu</i> 'peito'	<i>methe</i> 'querer'
/ɔ/	<o, O>	odo 'luto'	<i>gogo</i> 'gostar'	<i>po</i> 'arvore, pau'
/o/	<ô, Ô>	ôdô 'almofariz'	<i>gôgô</i> 'especie de arvore'	<i>po</i> 'poder'
/i/	<i, I>	<i>ligi</i> 'levantar'	<i>pidi</i> 'por causa de'	<i>bisi</i> 'vestir'
/u/	<u, U>	<i>dumu</i> 'pisar'	<i>fugudu</i> 'nervoso'	<i>vuvu</i> 'buraco, orificio'

Fonte: São Tomé e Príncipe (2013, p. 661)¹⁸

Sobre as vogais médias, o ALUSTP determina que:

Por apenas existirem os dois pares mínimos [e] - [ɛ] e [o] - [ɔ], estes são distinguidos mediante a utilização de acento circunflexo para marcar as médias fechadas [e] e [o], representadas como <ê> e <ô>, conforme ilustrado na Quadro 4. As médias abertas [ɛ] e [ɔ] não recebem qualquer acentuação. (STP, 2013, p. 662)

Com relação às vogais nasais¹⁹, estas deverão ser grafadas com <n> ou <m>, sendo grafadas com <m> a nasalidade que precede as consoantes bilabiais /p/ e /b/, conforme a Quadro 5 a seguir:

¹⁸ SÃO TOME E PRINCIPE. Decreto nº 19, de 14 de agosto de 2013. Dispõe sobre a ortografia e cria o Alfabeto Unificado para as Línguas Nativas de São Tomé e Príncipe. São Tomé, 2013.

¹⁹ As vogais orais longas que não sejam estilísticas são representadas por uma sequência de duas vogais idênticas. O alongamento vocálico é uma propriedade do Lung'le e do Ngola. Como esta tese centra-se no *Santome*, não foram incluídos exemplos de vogais orais longas.

Quadro 5 - Vogais nasais / nasalidade que precede /p/ e /b/

		Português
Santome	<i>Bembe</i>	agitar, balançar'
	<i>Tambôlô</i>	tambor
Ngola	<i>Kampu</i>	campo
	<i>Tambá</i>	pescar
Lung'le	<i>Kompwe</i>	compadre
	<i>Romba</i>	destruir

Fonte: São Tomé e Príncipe (2013, p. 662)²⁰

Nos demais contextos, a nasalidade é grafada com <n>.

Quadro 6 - Vogais nasais - nasalidade que precede grafemas diferentes de /p/ e /b/

Santome	Português
<i>San</i>	senhora
<i>Maxibin</i>	jovem
<i>Tamen</i>	tambor
<i>Mamon</i>	mamão
<i>Sun</i>	senhor
<i>Punta</i>	perguntar
<i>Mutêndê</i>	Palmeira pequena

²⁰ SÃO TOME E PRINCIPE. Decreto nº 19, de 14 de agosto de 2013. Dispõe sobre a ortografia e cria o Alfabeto Unificado para as Línguas Nativas de São Tomé e Príncipe. São Tomé, 2013.

<i>Kondê</i>	esconder
<i>Lung'le</i>	Português
<i>Kukundya</i>	coco
Kôndê	Esconder

Fonte: São Tomé e Príncipe (2013, p. 663)²¹

Quadro 7 - Consoantes segundo ALUSTP

Fonema	Grafemas	Santome	Lung'le	Ngola
/b/	<b, B>	<i>bêndê</i> 'vender', <i>ôbô</i> 'floresta'	<i>bala</i> 'bala' <i>pobe</i> 'pobre'	<i>pobe</i> 'pobre' <i>kubi</i> 'cobrir'
/d/	<d, D>	<i>dumu</i> 'pisar', <i>dôdô</i> 'maluco'	<i>daman</i> 'amigo', <i>mwadu</i> 'molhado'	<i>delu</i> 'dinheiro', <i>dooba</i> 'dobra'
/L/	<dh, DH>	-	-	<i>Dhuga</i> 'atirar', <i>Ridhê</i> 'dentro'
/dL L/	<dj, DJ>	<i>dja</i> 'dia', <i>djêlu</i> 'dinheiro'	-	<i>Djendja</i> 'depressa', <i>pandji</i> 'padrinho'
/f/	<f, F>	<i>safu</i> 'safu', <i>fla</i> 'falar', 'dizer'	<i>fê</i> 'feio', <i>sôfê</i> 'sofrer'	<i>fomi</i> 'fome', <i>fia</i> 'folha'
/g/	<g, G>	<i>ngê</i> 'pessoa', <i>bega</i> 'barriga'	<i>gô</i> 'chorar', <i>fega</i> 'esfregar'	<i>digô</i> 'fumo', <i>baga</i> 'desmanchar'
/gL b/	<gb, GB>	<i>gbêgbê</i> 'espécie de caracol'	<i>gbê</i> 'esmagar', <i>igbê</i> 'corpo'	-
/L /	<j, J>	<i>ja</i> 'dia', <i>pji</i> 'pedir'	<i>janela</i> 'janela', <i>vijin</i> 'vizinho'	<i>junhu</i> 'junho', <i>ji</i> 'fazer'
/k /	<k, K>	<i>kume</i> 'comer', <i>faka</i> 'faca'	<i>kudi</i> 'responder', <i>piku</i> 'pico'	<i>kathô</i> 'cão', <i>tuka</i> 'descer'
/kL p/	<kp, KP>	-	<i>ukperi</i> 'cesto', <i>ikpe</i> 'grão, semente'	-
/l/	<l, L>	<i>lembla</i> 'lembrar', <i>golo</i> 'procurar'	<i>lala</i> 'lá', <i>ulado</i> 'lado'	<i>longa</i> 'gamela', <i>thela</i> 'estrela'
/L /	<lh, LH>	<i>olha</i> 'orelha', <i>alha</i> 'areia'	<i>Familha</i> 'família', <i>milho</i> 'melhor'	-

²¹ SÃO TOME E PRINCIPE. Decreto nº 19, de 14 de agosto de 2013. Dispõe sobre a ortografia e cria o Alfabeto Unificado para as Línguas Nativas de São Tomé e Príncipe. São Tomé, 2013.

/m/	<m, M>	<i>mosu</i> 'rapaz' <i>sama</i> 'chamar'	<i>misa</i> 'missa' <i>umundu</i> 'mundo'	<i>maaku</i> 'macaco' <i>tometu</i> 'barulho'
/n/	<n, M>	<i>novu</i> 'novo' <i>anu</i> 'ano'	<i>nega</i> 'negar' <i>unu</i> "nu"	<i>neni</i> 'anel' <i>disinu</i> 'destino'
/L/	<nh, NH>	<i>nganha</i> 'galinha' <i>fanha</i> 'farinha'	<i>ginha</i> 'galinha', <i>nhan</i> 'sim, pois'	<i>nganha</i> 'galinha' <i>ngaa</i> 'lugar'
/p/	<p, P>	<i>poto</i> 'porta' <i>ope</i> 'pé, perna'	<i>pôkô</i> "porco" <i>pepelu</i> "papel"	<i>puga</i> 'purgar' <i>kapatazi</i> 'capataz'
/r/	<r, R>	<i>karu</i> 'carro'	<i>ranha</i> 'rainha' <i>aré</i> 'rei'	<i>ripa</i> 'limpar' <i>kuru</i> 'cru'
/s/	<s, S>	<i>sela</i> 'cheirar' <i>klason</i> 'calças'	<i>sopa</i> 'soprar', <i>posan</i> 'cidade'	<i>singô</i> 'pescoço' <i>ngosi</i> 'noite'
/t/	<t, T>	<i>tason</i> 'sentar' <i>kota</i> 'cortar'	<i>tuban</i> 'tubarão' <i>giitu</i> 'grito'	<i>tamba</i> 'pesca' <i>kata</i> 'cantar'
/L/	<th, TH>	-	-	<i>thentu</i> 'incenso' <i>potho</i> 'cidade'
/tL L/	<tx, TX>	<i>txila</i> 'tirar', <i>kitxiba</i> 'banana prata'	<i>txipa</i> 'intestino' <i>Idintxi</i> 'dente'	<i>txiba</i> 'banana' <i>katxi</i> 'meio'
/v/	<v, V>	<i>vôlô</i> 'zangar' <i>novu</i> 'novo'	<i>vozu</i> 'voz', <i>uventu</i> 'vento'	<i>vôrô</i> 'muito' <i>uvwa</i> 'nove'
/w/	<w, W>	<i>wo</i> 'agora' <i>awa</i> 'água'	<i>we</i> 'ir' <i>rêndêw</i> 'rendeiro'	<i>wê</i> 'olho' <i>awa</i> 'água'
/L/	<x, X>	<i>xê</i> 'sair', <i>naxi</i> 'ainda não'	<i>xinku</i> 'cinco' <i>maxi</i> 'mais'	<i>palaxu</i> 'palácio' <i>xiriga</i> 'mentira'
/j/	<y, Y>	<i>ya</i> 'eis' <i>saya</i> 'puxar'	<i>ya</i> 'eis, que', <i>uguya</i> 'agulha'	<i>yô</i> 'muito', <i>êya</i> 'sim'
/z/	<z, Z>	<i>zuda</i> 'ajudar' <i>oze</i> 'hoje'	<i>zuda</i> 'ajudar' <i>fooza</i> 'ferrugem'	<i>ziaru</i> 'branco', <i>zazi</i> 'trovão'

Fonte: São Tomé e Príncipe (2013, p. 664-665)²²

1.3.3. Língua Portuguesa em São Tomé e Príncipe

Português é a língua oficial de São Tomé e Príncipe e goza de um status diferente das demais línguas. Como língua trazida e imposta pelos colonizadores, sempre ocupou

²² SÃO TOME E PRINCIPE. Decreto nº 19, de 14 de agosto de 2013. Dispõe sobre a ortografia e cria o Alfabeto Unificado para as Línguas Nativas de São Tomé e Príncipe. São Tomé, 2013.

posição de prestígio, antes e depois da independência, na sociedade são-tomense de forma geral, enquanto as outras foram marginalizadas e até proibidas em algumas fases da colonização. Como língua de prestígio, era inicialmente reservada à elite são-tomense e promovida/estimulada pela classe dominante. E essa ‘política linguística’ adotada pelos colonizadores resultou no preconceito linguístico até mesmo entre os naturais de São Tomé e Príncipe em relação às línguas naturais faladas no arquipélago.

Assim sendo, eram considerados cultos e importantes aqueles que dominavam e falavam o “Português de Portugal” (PPT) e incultos e atrasados aqueles que não o dominavam. Esse fenômeno continua enraizado na sociedade são-tomense até hoje e fez com que as línguas nativas, não promovidas pelo Estado, não fossem incentivadas pelas famílias são-tomenses e, portanto, cada vez menos transmitidas e utilizadas de geração para geração no país.

Como consequência, as línguas nativas foram perdendo cada vez mais espaço entre os falantes e, como citado anteriormente, sendo mais comuns na população mais idosa e menos comuns entre os jovens, que representam de longe a maior porcentagem da população de STP atualmente. Em outras palavras, as línguas nativas de STP e seus falantes estão em pleno declínio e o português, por sua vez, com um número crescente de falantes. E é nesse contexto sócio-histórico que surgiu esse “português”, agora adequadamente chamado de Português de São Tomé (PST).

O PST foi e é completamente impactado pelas línguas dos povos africanos escravizados e levados para São Tomé durante a colonização. Na verdade, podemos afirmar que o PST é o resultado da convivência, ora conflituosa ora harmoniosa, durante séculos entre o PPT e as línguas nativas de São Tomé e Príncipe. Usando o português do Brasil como exemplo, Coelho (1881, p. 65) afirma que “...perdura na língua

portuguesa, não só na inflexão da voz, não só na phonetica, mas ainda no ternêo grammatical e no phraseado que tem *seu que* de novo não usado na terra lusitana, e a final em grande número de vocábulos de todo não portugueses”. E acrescenta:

E esses mamelucos, caboclos e caipiras, fallando a língua do “outro”, do estrangeiro, do homem de lá longe, do emboaba, (*amo abá*), fallando essa língua corrompida pelo fallar do africano, do selvagem negro (*tapyyñuna*), conservam no sotaque, no phraseado, reminiscencias da “língua geral” que vão se-fazer ouvir ainda no seio do parlamento, onde desgraçadamente predomina um francez assaz eivado de francezismos e tambem já do não poucos inglezismos. (Coelho, 1881, p. 65)

Em São Tomé e Príncipe, acontece(u) um fenômeno semelhante ao que Coelho (1881) cita a seguir e o PST, objeto de estudo desta pesquisa, apresenta inúmeros casos lexicais, terminológicos e sintáticos que serão devidamente analisados e discutidos mais adiante, no Capítulo 4 desta tese. Nesse quesito, firme no seu posicionamento, o pesquisador reforça o entendimento de Coelho (1881) com o argumento que segue:

Assim como os selvagens ou desapareceram ou mestiçou-se no rústico fallar do nosso povo, conseguindo introduzir na lingua portugueza do Brazil centenaes de raizes. O cruzamento d'estas raças (indigenas do Brazil, negros e brancos) ao passo que misturou os sangues, cruzou tambem (se nos é lícito servirmo-nos d'essa expressão) a lingua, sobretudo a linguagem popular. É assim que, na linguagem do povo das provincias do Pará, Goyaz, e especialmente de Matto-Grosso, há não só quantidade de vocabulos tupis e guaranis accomulados á lingua portugueza e n'ella transformados, como ha phrases, figuras, idiotismos, e construcções peculiares ao tupi. Este facto mostra que o cruzamento physico de duas raças deixa vestigios moraes, não menos importantes do que os do sangue. (Coelho, 1881, p. 66)

Conforme apresentado no Capítulo 1, administrativamente, São Tomé e Príncipe é dividido em sete distritos (Figura 10, p. 44), sendo o Pagué²³ hoje a Região Autônoma

²³ As grafias das regiões administrativas/distritos, topônimos, jornal, revista e o programa de TV foram mantidas nesta tese por serem as designações oficiais, mas apresentamos aqui as grafias segundo o ALUSTP: *Pagué* → *Page*, *Mé-Zochi* → *Me-Zoxi*, *Caué* → *Kawe*, *Ió* → *Yô*, *Téla Nón* → *Tela Non*, *Muala Glavi* → *Mwala Glavi*, *Klonvesá* → *Klonvesa*

do Príncipe (RAP). A designação dos distritos que compõem o arquipélago de São Tomé e Príncipe, por sua vez, é baseada especialmente nos nomes dos principais rios que os cortam: Água Grande (Cidade de São Tomé, capital), Mé-Zochi (Cidade de Trindade), Cantagalo (Cidade de Santana), Lembá (Cidade de Neves), Lobata (Cidade de Guadalupe), Caué (Cidade de Angolares), em São Tomé, e Pagué, na Região Autónoma do Príncipe (Cidade de Santo António). Podemos observar que cinco dos sete distritos parecem ter sido nomeados com palavras das línguas africanas trazidas pelos trabalhadores escravizados e/ou das línguas nativas de STP – Mé-Zochi, Lembá, Lobata, Caué e Pagué. Além desses, existem diversos outros rios, ilhéus, praias cujos nomes remetem à ancestralidade africana: Rio Ió Grande, Praia Ió Grande, Ponta Ió Macaco, Ilhéu Quixibá etc. Portanto, vemos aqui a importância da toponímia para estudos linguísticos em geral e para o caso de São Tomé e Príncipe em particular.

Além desses, o Santome se faz (ou se fez) presente no PST também nos nomes de partidos ou coligações político-partidárias - *Códó* (Coligação de S. Tomé e Príncipe - [Corda, coligação, união]) e *Uê Kédáji* [Caminho da Luz], nas eleições legislativas de 1991 e 2002, respectivamente (ver Quadro 8, p. 59). Por fim, e não menos importantes, temos e tivemos jornais – *Téla Nón*, *Revista Muala Glavi*, programa de TV – *Klonvesá*, lema de campanha política (Figura 12, p. 60) etc.

Quadro 8 - Partidos políticos na língua Santome

PST/ST (Designação oficial)	ALUSTP	Português	Sentido
Códó	<i>Kodo</i>	Corda	União
Uê Kedagi	<i>Wê Kedaji</i>	Olho de claridade	Luz, Esperança
Bamu Kêlê (lema do MLSTP-PSD)	=	Vamos acreditar	Fé, Esperança

Fonte: Elaborado pelo autor

Figura 12 – Slogan do Partido MLSTP nas Eleições legislativas de 2022



Fonte: Jorge Bom Jesus/MLSTP (2022)²⁴

1.4. SÍNTESE DO CAPÍTULO

O objetivo deste capítulo foi apresentar a história geral de São Tomé e Príncipe, desde a chegada dos colonizadores portugueses nos finais do século XV, passando pelos processos de povoamento, a exploração econômica por meio dos grandes ciclos com introdução das diferentes culturas/plantações, cana-de-açúcar, cacau e café. É também neste capítulo que discutimos a polêmica envolvendo a origem dos Angolares e

²⁴ Disponível em: <https://www.facebook.com/share/p/zA2yUMm2mXTofbix/?mibextid=oFDknk>. Acesso em: 21 mai. 2024.

do Rei Amador, a resistência e lutas dos são-tomenses pela independência, o processo de descolonização, independência, a atual organização política, situação socioeconômica, cultural e, principalmente, o cenário linguístico do país. Demonstramos, por meio de dados, que essa situação linguística de STP, centro desta pesquisa, é o resultado da sua história e das dinâmicas sociais pelas quais as ilhas passaram e passam. Assim sendo, discutiremos no próximo capítulo as abordagens teóricas que sustentam os processos de mudança linguística por meio da lexicalização e da gramaticalização do PST, que também têm o contato linguístico na sua gênese.

CAPÍTULO 2: PERSPECTIVAS TEÓRICAS

Neste capítulo, apresentamos uma sucinta discussão dos estudos linguísticos contextualizados relacionados à lexicalização e à gramaticalização bem como *code-switching*. Assim sendo, antes das considerações de lexicalização, gramaticalização e *code-switching*, faz-se necessário tratarmos primeiramente das questões de fundo - a linguística em si. Os três fenômenos em questão são manifestações de tensões internas inerentes a dinâmicas da própria língua e pressões externas resultantes do uso e do contato com outras línguas, em vista dos seus usuários e contextos de uso e situações comunicativas. Lexicalização e gramaticalização podem revelar indícios de mudança linguística em curso ou concluída, e o *code-switching*, um fator que poderá ser um motivador ou o próprio gatilho dessa mudança. Portanto, nesta pesquisa, seguiremos uma abordagem funcionalista da mudança linguística.

Para Maria Helena de Moura Neves, funcionalismo ou gramática funcional é “uma teoria da organização gramatical das línguas naturais que procura integrar-se em uma teoria global da interação social.” (Neves, 2021, p. 15) Ainda assim, a autora ressalta o desafio de caracterizar o Funcionalismo uma vez que:

Os rótulos que se conferem aos estudos ditos “funcionalistas” mais representativos ligam-se diretamente aos estudiosos que os desenvolveram, não propriamente a características definidoras da corrente teórica em que eles se colocam. No final do século XX, Prideaux (1994) afirmava que provavelmente existiriam tantas versões do Funcionalismo quantos linguistas denominados “funcionalistas”, um termo que abrange desde os que simplesmente rejeitam o Formalismo até os que criam uma teoria a que atribuem essa filiação. O que se pode afirmar e defender, na verdade, é que, dentro do que vem sendo denominado – ou autodenominado – “Funcionalismo”, encontram-se modelos teóricos muito diferentes. (Neves, 2021, p. 15)

Anteriormente, Dik (1997) preferiu contrapor as duas correntes através de paradigmas – formal e funcional. Paradigma é proposto para designar cada conjunto de crenças e hipóteses em interação. Segundo Dik (1997 *apud* Neves, 2021, p. 56), “enquanto o paradigma formal (PFO) vê a língua natural como um objeto abstrato, um conjunto de orações, e a gramática como uma tentativa para caracterizar esse objeto formal em termos de regras de sintaxe formal, no paradigma funcional (PFU), língua natural é vista como um instrumento de interação social; portanto, não existindo por si nem em si só como uma estrutura arbitrária de alguma espécie, mas em virtude de ser usada para certos propósitos, os quais concernem à interação entre os seres humanos”. Assim sendo, a função da língua para PFO é a expressão dos pensamentos enquanto para o PFU, é o estabelecimento de comunicação entre os usuários. Comunicação é, para Neves (2021):

Um padrão interativo dinâmico de atividades por meio das quais os usuários efetuam mudanças na informação pragmática de seus parceiros. A comunicação, assim, não é restrita à transmissão e à recepção de informação factual. Daí resulta que o uso da linguagem exige pelo menos dois participantes, o falante (*the speaker*, S) e o destinatário (*the addressee*, A).” (Dik, 1997 *apud* Neves, 2021, pp. 56-57)

Quadro 9 - Paradigma Formal vs Paradigma Funcional

TÓPICOS	Paradigma Formal	Paradigma Funcional
Língua	Conjunto de orações	Instrumento de interação social
Principal função da língua	Expressão de pensamentos	Comunicação
Correlato psicológico	Competência: capacidade de produzir, interpretar e julgar orações	Competência comunicativa: habilidade de interagir socialmente com a língua.
Sistema e uso	Estudo da competência tem prioridade sobre o da atuação.	Estudo do sistema deve fazer-se dentro do quadro do uso.
Língua e contexto/situação	As orações da língua devem	A descrição das expressões

	descrever-se independentemente do contexto/situação	deve fornecer dados para a descrição de seu funcionamento num dado contexto.
Aquisição da linguagem	Faz-se uso de propriedades inatas, com base em um input restrito e não estruturado de dados.	Faz-se com a ajuda de um input extenso e estruturado de dados apresentado no contexto natural.
Universais linguísticos	São propriedades inatas do organismo humano	Explicam-se em função de restrições: comunicativas; biológicas ou psicológicas; contextuais.
Relação entre a sintaxe, a semântica e a pragmática	A sintaxe é autônoma em relação à semântica; as duas são autônomas em relação à pragmática; as prioridades vão da sintaxe à pragmática, por via da semântica.	A pragmática é o quadro dentro do qual a semântica e a sintaxe devem ser estudadas; as prioridades vão da pragmática à sintaxe, por via da semântica.

Fonte: Adaptado de Neves (2021, p. 59)

Segundo Neves (2021, p. 15) não é fácil caracterizar Funcionalismo, ou Gramática funcional, uma vez que vários estudiosos com diversas abordagens se autoidentificam como funcionalistas, mas a autora define a corrente como a “teoria da organização das línguas naturais que procura integrar-se em uma teoria global da interação social”. E complementa: “Trata-se de uma teoria que, propondo que as relações entre as unidades e as funções das unidades têm prioridade sobre seus limites e sua posição, entende a gramática como suscetível às pressões do uso.”

Citando Gebruers (1987, p. 29), Neves afirma que:

O que caracteriza a concepção de linguagem defendida pela Gramática funcional - refletindo o defendido pela Escola Linguística de Praga - é o seu caráter não apenas funcional, mas também dinâmico. Ela é funcional porque não separa o sistema linguístico, e suas peças, das funções que têm de preencher, e é dinâmica porque reconhece, na instabilidade da relação entre a estrutura e função, a força dinâmica que está por detrás do constante desenvolvimento da linguagem.” (Gebruers, 1987 *apud* Neves, 2021, p. 18)

No âmbito funcional, Neves (2021, pp. 15-16) se apoia em Mackenzie (2016b) para afirmar que “As línguas têm uma função primordial nas sociedades humanas, a de permitir uma eficiente comunicação de ideias e de sentimentos entre os homens, função impacta tanto o uso corrente das línguas quanto o modo como elas se desenvolvem historicamente”.

Além das características, os funcionalistas também apresentam diferenças em diversos aspectos:

O exato objeto de pesquisa (discurso, texto, estruturas gramaticais); o tipo de dados que se admitem (dados de corpus, dados experimentais, julgamentos intuitivos, exemplos de gramáticas); o grau em que a língua é vista como sistema; o grau de entrosamento com outras subdisciplinas da Linguística (em especial a Linguística Cognitiva, a Gramática de construções, a Tipologia linguística), ou mesmo com outras disciplinas, como a Sociologia, a Psicologia, a Antropologia, a Pedagogia. (Mackenzie, 2016b *apud* Neves, 2021, p. 16)

Já em 1994, Martinet (p. 14) observa que o linguista deve ser constantemente guiado pela competência comunicativa porque toda língua se impõe tanto em seu funcionamento quanto em sua evolução, como instrumento de comunicação da experiência, podendo ser o que o ser humano sente, percebe, entende em todos os momentos da sua vida.

Assim, no Funcionalismo, nesta tese, *função* é usada segundo Halliday (1973a, p. 104) “papel que a linguagem desempenha na vida dos indivíduos, servindo a muitos e variados tipos universais de demandas, não se refere aos papéis que as classes de palavras ou sintagmas desempenham dentro da estrutura das unidades maiores”. (Halliday, 1973a *apud* Neves, 2021, p. 21). Porém, apoiando-se em Beaugrande (1993, cap. III, p. 3), Neves (1997) afirma que “A principal tarefa de uma “gramática funcional” é “fazer correlações ricas entre forma e significado dentro do contexto global do discurso”.

E para discutir as peculiaridades desta pesquisa, as reflexões e aplicações da Linguística Centrada no Uso (LCU) ou Teoria Centrada no Uso se enquadra muito bem, já que este trabalho foi motivado pelo uso do Português de São Tomé (PST), considerando seus usuários e contextos de uso. Sobre Linguística Centrada no uso [*Usage-based Linguistics*], Neves (2021, p. 129) ressalta:

Uma ligação íntima entre a estrutura linguística e seu uso linguístico, entendendo-se que as estruturas linguísticas estão atreladas aos eventos de uso, tanto em termos de memória e de aprendizagem quanto em termos de processamento (a produção e a compreensão). A análise da estrutura linguística desenvolve-se, pois, a partir dos processos cognitivos, alguns diretamente ligados com o uso da língua (por exemplo, a categorização), outros aparentemente não tão ligados (por exemplo, a memória). Essa centração no uso significa entender-se que os padrões da língua são generalizações obtidas pela repetição dos usos em interações (produções e compreensões), formando redes similares compartilhadas que constroem o sistema linguístico dos falantes: repetidas no uso, as estruturas linguísticas moldam representações cognitivas (rotinas cognitivas aprendidas), que, por sua vez, resultam em padrões de uso. (Neves, 2021, p. 129)

E Martelotta (2011, p. 53) argumenta que:

Embora o fenômeno de variação tenha influência, de fato, sobre o desenvolvimento de novos valores para as estruturas da língua, o que parece impulsionar o movimento de mudança são os mecanismos interacionais associados aos usos dessas estruturas. Em termos mais específicos, a mudança tende a refletir o modo mais eficaz de negociação do sentido que o falante e ouvinte promovem no ato da comunicação. (Martelotta, 2011, p. 53)

Porém, a autora esclarece que a LCU não constitui uma linha teórica definida, representando, na prática, uma conjunção de diferentes teorias da Linguística e reforça que a proposta centrada no uso não é um campo restrito à Linguística cognitiva, mas uma “agenda de pesquisa” que congrega pesquisadores também da Linguística Funcional (Givón, 2010) e da Psicolinguística (Bates, 1976) (Neves, 2021, p. 130).

Para Castilho (2011, p. 71), “A ciência clássica funciona bem na descrição e história de produtos cristalizados. O problema é que tanto a descrição da oralidade quanto a investigação diacrônica focalizam maiormente os processos linguísticos. E o autor reforça:

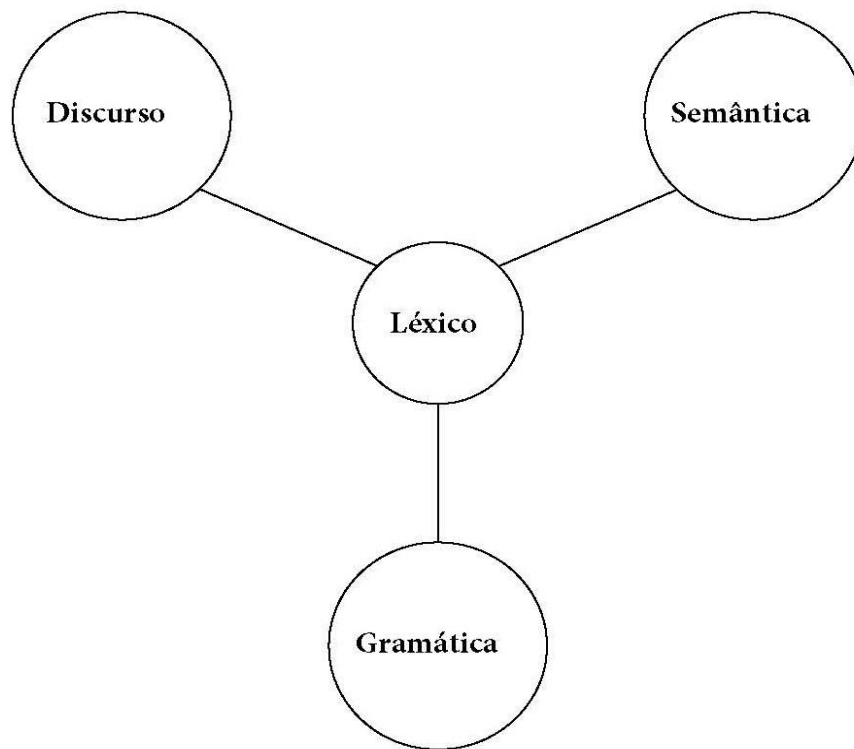
Para adequar a teoria a essa empiria, precisaremos operar com a ciência dos sistemas complexos. As seguintes afirmações configuram os sistemas complexos: (1) Os componentes dos sistemas complexos exibem um tipo de ordem sem periodicidade, em fluxo contínuo, em mudança – como queria Heráclito; (2) Os sistemas não são lineares, são dinâmicos, exibem um comportamento irregular, imprevisível. (3) Os elementos dos sistemas complexos exibem relacionamentos simultâneos, não são construídos passo a passo, linearmente. Eles são adaptáveis e auto-organizados. (4) As anomalias identificadas pela abordagem clássica exemplificam fenômenos vitais para o entendimento do problema, e não deveriam ser descartadas como aberrantes. (5) Uma nova topologia do impreciso, do vago, do aproximativo, precisará ser proposta. (6) A competição nos sistemas é mais importante que sua consistência. (7) Finalmente, ao tratar de fenômenos complexos, nenhum método revelará por si mesmo o objeto por inteiro. (Castilho, 2011, p. 71)

E dado a complexidade do ambiente linguístico-cultural em que se insere o estudo, consideraremos inicialmente língua, conforme Castilho (2013, p. 3-5) quando afirma que: “língua é um multissistema dinâmico, que pode ser graficamente representado numa forma radial, tendo ao centro o Léxico e à volta a Semântica, o Discurso e a Gramática”; “gramaticalização como é um processo dentre outros processos de criação linguística” e a “lexicalização como processo de criação de itens, dispostos com maior ou menor clareza nas classes de palavra ou categorias lexicais”, em que o léxico é “um conjunto de categorias cognitivas prévias à enunciação, com base nas quais construímos os traços semânticos inerentes”.

Nessa postulação teórica multissistêmica, Castilho (Figura 13, p. 68) argumenta que os sistemas são independentes uns dos outros (apesar do léxico estar no centro), dispondo cada um de categorias próprias, admitindo “que nossa mente opera simultaneamente sobre o conjunto das categorias recolhidas nesses sistemas – as

categorias lexicais, discursivas, semânticas e gramaticais, e defende que qualquer expressão linguística exhibe simultaneamente propriedades lexicais, discursivas, semânticas e gramaticais.

Figura 13 – Multissistema Linguístico de Castilho



Fonte: Castilho (2013, p. 4)

Um exemplo que se enquadra perfeitamente nessa discussão é o item lexical do PST *uêchaismo*, que pode ser traduzido como “inveja” e será objeto de análise mais detalhada no Capítulo 4 (p. 100). A raiz desse item lexical é constituída pela justaposição de duas palavras do Santome *uê* [olho] e *cha* [cheio] à qual se acrescentou o sufixo “ismo” do português, e deu origem ao “uechaismo”.

Verificamos nesse item lexical, basicamente, todas as características apontadas por Castilho: dinamicidade, comportamento irregular e imprevisível, anômalo e impreciso. Além disso, exibe simultaneamente propriedades lexicais, discursivas, semânticas e gramaticais, embora o léxico seja o núcleo.

2.1. LEXICALIZAÇÃO

Por uma questão metodológica e de organização desta tese, apresentaremos, de modo sucinto, as discussões teóricas dos subitens lexicalização e gramaticalização separadamente. No entanto, dada à proximidade da relação entre os dois processos, com fronteiras não muito bem delimitadas e discussão teórica apresentada quase sempre de forma comparativa, ambos os processos aparecerão repetidas vezes nas discussões dos dois subitens. Todavia, antes da discussão sobre lexicalização propriamente dita, teceremos algumas considerações sobre o conceito de palavra. Para tanto, recorreremos à Margarida Basílio e no livro “Teoria Lexical” de 1987. Sobre palavra, Basílio (1987) afirma:

A palavra é uma dessas unidades linguísticas que são muito fáceis de reconhecer, mas bastante difíceis de definir, se tomarmos como base de definição a língua falada. Isto acontece porque na língua falada não fazemos pausas sistemáticas entre cada palavra pronunciada. Na língua escrita, não temos problemas de definição neste eixo, e podemos definir palavra como qualquer sequência que ocorra entre espaços e/ou sinais de pontuação, estamos nos referindo, naturalmente, a sequências possíveis na língua; uma sequência como *sqwrn* seria interpretada provavelmente como um erro de datilografia, nunca como uma palavra do português. (Basílio, 1987, p. 6)

Naturalmente que a sequência *sqwrn*, dada como exemplo por Basílio, antecede a criação da World Wide Web e das redes sociais e, conseqüentemente, dos internetês

(linguagem dos internautas) entre os quais *blz* (beleza), *sqn* (só que não), *sdds* (saudades), *rsrs* (risos), *tbm* (também), *kd* (cadê), *vlw* (valeu), *tmj* (tamo junto) etc. e demandaria discussões e análises bem mais aprofundadas sobre possíveis comparações e analogias. A menção aqui se faz necessária uma vez que esta tese utiliza como base a oralidade e os dados analisados terem sido extraídos essencialmente das mídias sociais.

Basílio ressalta, no entanto, que há outro eixo em que a definição de palavra causa dificuldades – a distinção que normalmente se estabelece entre duas palavras distintas e duas formas da mesma palavra. A autora apresenta *casa* e *casas* tidas como exemplos de duas formas da mesma palavra e *casa* e *casinhola*, entendidas como duas palavras diferentes. Além disso, afirma também que “dizemos que *falam* e *falamos* são duas formas do verbo falar, mas que *falante* constitui uma palavra diferente”, e salienta que geralmente a diferença entre palavras distintas e diferentes formas da mesma palavra resultam da diferença entre flexão e derivação. Nesse sentido, levanta a questão de não haver uma distinção nítida e definitiva entre os conceitos de flexão e derivação, afirmando que “muitas vezes a diferença entre os dois conceitos é colocada como decorrente da diferença entre “palavras distintas” e “formas de uma mesma palavra”. Forma-se, portanto, um círculo vicioso”.

Após essas breves considerações, retomemos a discussão sobre lexicalização. Moreno Cabrera (1998, p. 214) define lexicalização como o processo de criação de itens lexicais a partir de unidades sintáticas. Tomando expressões idiomáticas como exemplo, o autor as define como “construções sintáticas que perdem a sua composicionalidade (propriedade das expressões complexas cujo sentido é determinado pelos sentidos dos seus constituintes) e adquirem um novo conteúdo idiossincrático. Esses sentidos se

lexicalizam e podem ser registrados no léxico como todo não-analisável, o que o autor chama de “processo lexicotético”. Assim sendo, argumenta que a lexicalização vai exatamente no sentido oposto da gramaticalização e, portanto, os processos de lexicalização deveriam seguir a mesma hierarquia da gramaticalização, mas no sentido contrário.

Moreno Cabrera (1998, pp. 223-225) afirma também que o princípio da unidirecionalidade geralmente atribuída apenas à gramaticalização não é uma propriedade exclusiva desta e, portanto, deveria ser considerado no estudo da evolução e mudança linguística como um todo e propõe o uso do termo “irreversível” no lugar de “unidirecionalidade”. O autor argumenta que se fosse exclusivo da gramaticalização, as línguas deveriam se tornar cada vez mais gramaticalizadas e isso não acontece, pois a evolução linguística é um processo bidirecional e contempla tanto a gramaticalização quanto a lexicalização.

Moreno Cabrera ressalta ainda que, no processo de mudança linguística, existe movimento constante do léxico para a sintaxe e vice-versa e não se observa as línguas perdendo o seu léxico e enriquecendo a morfologia e sintaxe, nem aumentando gradualmente seu léxico e reduzindo sua morfologia e sintaxe. Portanto, a evolução linguística não resulta apenas do processo gramaticalização nem de lexicalização; mas é a interação entre ambos que pode produzir resultados equilibrados que observamos na evolução das línguas, ressalta. Por fim, o autor afirma que os processos de gramaticalização são predominantemente metafóricos e os processos de lexicalização são majoritariamente metonímicos.

Por sua vez, para DeLancey (2001, p. 102), ao perder a sua autonomia, um morfema lexical pode seguir a uma das trajetórias, ou a gramaticalização ou a

lexicalização. O autor define lexicalização como o processo pelo qual lexemas autônomos se tornam parte de novos itens lexicais. Para DeLancey, o termo lexicalização também é usado para descrever o processo oposto à gramaticalização, em que uma forma flexionada deixa o seu paradigma e se torna um lexema distinto, como na formação de advérbios a partir de formas de caso oblíquo. Ainda argumenta que a principal distinção entre lexicalização e gramaticalização está no resultado do processo: a lexicalização resulta em um novo lexema, enquanto a gramaticalização resulta em uma nova construção gramatical; adverte ainda que essa distinção não é uma dicotomia rígida, mas um *continuum*, dada a natureza da fronteira entre morfemas lexicais e gramaticais.

Ao apresentar, a origem do prefixo *a-* em *atop* (do inglês, 'no topo') como um exemplo de lexicalização, DeLancey afirma que, nesse caso, a redução fonológica e a cliticização de uma preposição resultaram em um único item lexical, não numa nova construção. Como contraponto, considera que o processo de *top* (topo) sair da classe aberta de substantivos para se tornar um membro da classe fechada de preposições seria, contudo, mais adequadamente discutido como gramaticalização. DeLancey lista outros casos como os adjetivos que começam com *a-* ("*awake*", "*asleep*" e "*alone*") que representam uma nova subcategoria lexical, mas as formas resultantes diferem da maioria dos adjetivos por não poderem ocorrer em posição pré-nominal (ex: **an alone man*). Nesse caso, ele enfatiza que, embora crie uma nova subcategoria, não há uma nova construção gramatical associada, e por isso não seria útil classificá-la como gramaticalização. Em outras palavras, a descategorização de um lexema não ocorre sem uma mudança funcional prévia em direção à gramaticalização ou à lexicalização. Para ser considerado gramaticalização, o critério mais propício seria o grau em que o

resultado do processo é uma nova construção produtiva, isto é, um novo elemento da estrutura gramatical, em oposição a uma nova forma lexical ou um novo conjunto de formas lexicais – lexicalização.

Nesta tese, porém, seguiremos essencialmente na direção do conceito de lexicalização apresentado por Castilho. Segundo Fortunato (2008, p. 1396), Castilho (2003b) afirma que:

A **Lexicalização** seria o processo de criação das palavras por seleção de categorias cognitivas e de traços semânticos derivados, processando-se sua concentração num dado item, o qual é composto por um conteúdo semântico e uma sequência fonológica. Uma vez criadas, as palavras passam por alterações em suas categorias e subcategorias cognitivas, tanto quanto em seus papéis semânticos. (Castilho, 2003b *apud* Fortunato, 2008, p. 1396) (grifo no original).

Podemos afirmar que essa definição de lexicalização proposta por Castilho se enquadra de forma relevante nas discussões de sistema linguístico e o seu uso, trazidas por Simon Dik. Dik (1989, p. 6) argumenta que as línguas naturais são um instrumento utilizado para fins de comunicação e não faz muito sentido analisar as propriedades sem considerar os usos e funções. O sistema subjacente à construção de expressões linguísticas é um sistema funcional e deve sempre ser estudado observando as regras, princípios e estratégias que determinam seu uso comunicativo natural. Para enfatizar, Dik (1989, p. 6) argumenta:

Em outras palavras, a questão de como uma língua é organizada não pode ser estudada de forma adequada sem levar em conta a questão do porquê ela é organizada da forma como é, dadas as funções comunicativas que cumpre. Isso significa que expressões linguísticas só podem ser compreendidas adequadamente quando são consideradas como funcionando em contextos cujas propriedades são co-determinadas pelas informações contextuais e situacionais disponíveis para os usuários e os seus interlocutores. A língua não funciona de forma isolada: ela integra uma realidade humana (psicológica e social) viva. (Dik, 1989, p. 6) (tradução nossa)

2.2. GRAMATICALIZAÇÃO

O primeiro a utilizar o termo gramaticalização foi Meillet (1912, p. 131), para quem gramaticalização “*consiste dans le passage d’un mot autonome au rôle d’element grammatical*” [passagem de uma palavra autônoma ao papel de elemento gramatical] (tradução nossa). Nesse sentido, segundo Neves (2021, p. 165), Meillet afirma que “a gramaticalização de certas palavras cria novas formas, introduz categorias que não tinham expressão linguística, e, assim, transforma o sistema como um todo. Portanto,

O processo de gramaticalização propõe-se, pois, como, essencialmente, uma passagem do ‘lexical’ ao ‘gramatical’, envolvendo, nesta ponta, uma sequência interna de sintático para morfológico. Isso representa um *continuum* que, afinal, pode ser representado como ‘lexical > sintático > morfológico’, fórmula que permanece nos trabalhos subsequentes sobre esse processo denominado (no geral das apresentações) como gramaticalização. (Neves, 2021, p. 165)

Lehmann (1995 [1982]), por sua vez, define a gramaticalização como um processo que transforma lexemas em formativos gramaticais e formativos gramaticais em mais gramaticais ainda (Gonçalves et al., 2007, p. 70). Também citando Lehmann (2002), Neves (2021, p. 166) afirma:

[A gramaticalização e a lexicalização] dois processos - aparentemente opostos - são complementares: a gramaticalização envolve acesso analítico a uma unidade (a análise das partes da unidade seguida da análise de seu todo), enquanto a lexicalização envolve acesso holístico a uma unidade (a análise de uma unidade como um todo, sem consideração de análise interna). Assim, a gramaticalização ocorre quando um signo adquire funções na formação analítica de signos mais abrangentes, enquanto a lexicalização ocorre quando um signo, retirado do acesso analítico, passa a fazer parte do inventário lexical da língua. Os dois processos não dizem respeito de forma isolada a sinais, mas referem-se a uma sinalização nas relações paradigmáticas e sintagmáticas; e, na verdade, ambos os processos ocorrem na direção comunicativa de prover o falante dos meios necessários de expressão. (Lehmann, 2002 *apud* Neves, 2021, p. 166)

Heine et al. (1991b, pp. 5-11) fizeram um levantamento histórico de pesquisa sobre gramaticalização, que começa no século X na China, no séc. XVIII França (Condillac e Rousseau) e na Inglaterra (Tooke), depois no séc.XIX na Alemanha (com Bopp, Schlegel, Humboldt, Gabelentz) e Whitney nos Estados Unidos. E, embora Meillet tenha sido o primeiro a utilizar o termo gramaticalização, Heine et al. (1991b, p. 5) afirmam:

Foi J. Home Tooke, contemporâneo de Condillac, que pode ser considerado o pai dos estudos sobre gramaticalização. Para Tooke, o "segredo" das palavras está na sua etimologia. Uma noção fundamental no seu trabalho, tornada pública pela primeira vez em 1786 (e depois em 1805) foi publicada em volume único (Horne Tooke 1857), é "abreviação": nomes e verbos são chamados de "palavras necessárias" e são considerados partes essenciais do discurso, enquanto as outras classes de palavras, tais como advérbios, preposições, e conjunções, resultam da abreviação ou "mutilação" das "palavras necessárias. As formas inflexionais e derivacionais são tratadas por ele como fragmentos de palavras antes independentes aglutinadas ao radical (cf. Robins [1967] 1979:155-58). (Heine et al., 1991b, p. 5) (tradução nossa)

Heine et al. (1991b) apontam Jerzy Kurylowicz (contemporâneo de Meillet) para quem a gramaticalização era um processo de morfologização, tendência adotada por estudiosos modernos até os mais atuais, “onde o foco passou a ser quase exclusivamente sincrônico, com mudança linguística sendo vista como o ajustamento entre estágios sincrônicos isolados”. Kurylowicz (1975 [1965], p. 52) definiu gramaticalização como o “avanço de um morfema do estatuto lexical para o estatuto gramatical, ou de um estatuto menos gramatical para um estatuto mais gramatical, por exemplo, do estatuto formante derivativo para o de um formante flexional”.

Neves (2021, p. 169) apresenta igualmente uma definição trazida por Hopper & Tragott (1993: xv) para quem “gramaticalização é definida como o processo pelo qual, em determinados contextos linguísticos, itens e construções passam a servir a funções gramaticais, e, uma vez gramaticalizados, continuam a desenvolver novas funções

gramaticais. No entanto, Neves (2021, p. 168) atribui a Givón a revitalização do estudo do tema com seu *slogan* “A morfologia de hoje é a sintaxe de amanhã” em 1971, que demonstrou, com a evidência de línguas africanas, que formas verbais que hoje se constituem de radicais com afixos remontam a arranjos de pronomes com verbos independentes e a proposição do ciclo “discurso > sintaxe > morfologia > morfofonêmica > zero” e cunhou o termo sintetização.

Coelho (2010, pp. 335-336), por seu turno, define gramaticalização como um “processo segundo o qual um item tramita de categoria na língua, destituindo-se de funções referenciais, para assumir funções gramaticais” ou também “quando um item se desloca dentro de uma mesma categoria, acentuando apenas seu grau de abstração e assumindo, paulatinamente, maior capacidade de expressar funções gramaticais”. Por sua vez, Coelho (2010, p. 338) define lexicalização ou expansão lexical como o “processo de expansão de semas de um item decorrente de sua abstração que rumo à categoria gramatical e que lhe permite aumentar as funções de referenciação no âmbito do léxico. A autora argumenta ainda que,

No curso diacrônico da gramaticalização, um item linguístico também pode perder alguns de seus antigos semas. Isso decorre de fatores tanto linguísticos quanto extralinguísticos. A esse processo de esvaziamento semântico por que se passam os itens linguísticos denominou-se de deslexicalização ou restrição semântica. (Coelho, 2010, p. 340)

Em síntese, Coelho (2010, p. 341) defende que gramaticalização e lexicalização são fenômenos linguísticos que se processam num *continuum* paralelo porque se expandem em categorias linguísticas distintas, mas que, de certa forma, relacionam-se, uma vez que a gramaticalização é um subconjunto oriundo da polissemia.

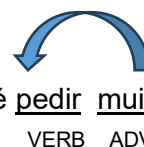
Apoiando-se em Castilho (2006), Gonçalves et al. (2007, p. 16) afirmam que a gramaticalização ocorre mediante três subprocessos: fonologização (alterações no corpo fônico das palavras), morfologização (alterações que afetam o radical e os afixos) e sintaticização (alterações que afetam as categorias lexicais, os arranjos sintagmáticos e a atribuição de funções na sentença). Camara Jr (1986 *apud* Fortunato, 2008, p. 1394) denomina de palavras lexicais e as define como “vocábulo providos de semantema, ou seja, têm traços que remetem ao ambiente bio-social extralinguístico”. É nesse contexto que Hopper (1987) afirma:

A gramática das línguas é constituída de partes cujo estatuto vai sendo constantemente negociado na fala, não podendo em princípio ser separado das estratégias de construção do discurso. Subjazem a esse entendimento uma concepção de língua como atividade no tempo real e a postulação de que, a rigor, não há gramática como produto acabado, mas sim em constante gramaticalização. (Hopper, 1987 *apud* Gonçalves et al., 2007, p. 15)

Embora existam diferentes termos em uso, nesta tese adotamos gramaticalização na concepção de Gonçalves et al. (2007, p. 17): “alterações de propriedades sintáticas, semânticas e discursivo-pragmáticas de uma unidade linguística que promovem a alteração de seu estatuto categorial”. Os exemplos 11b e 11c, a seguir, ilustram esses casos:

(11) **Quantificadores e intensificadores no PST**

- a) Só quero uma relação que permite dar 4 round por dia, é pedir muito?



O item lexical “muito” não pertence a uma categoria gramatical fixa e desempenha diferentes funções em diversos contextos de uso, podendo se comportar como um

advérbio ou pronome indefinido. Na alínea 11a, “muito” é invariável e intensifica ação do verbo “pedir”. Funciona, portanto, como advérbio de intensidade.

b) **Aqui tá com ladrão muito.**
 NOM PRON IND ADJ

*Naí sa ku ladlon muntu.*ST

Aqui está com muito ladrão.^{PST/PT}

c) **Ladrão muito tá n São Tomé.**
 NOM PRON IND ADJ VERB PREP NOME

*Ladlon muntu sa (ni) Santome.*ST
 LADRÃO PRON IND ADJ ESTAR (PREP) SÃO TOMÉ

São Tomé está com muitos ladrões. / São Tomé está cheio de ladrões.^{PST/PT}

d) **San se té wêxa muntu.**^{ST25}
 SENHORA DEM TER OLHO CHEIO ADV

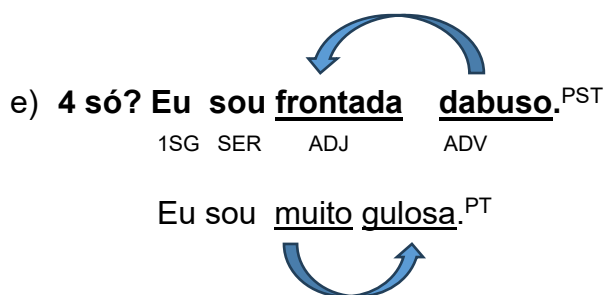
Essa Senhora tem olho cheio muito.^{PST}

Essa Senhora é muito invejosa.^{PST/PT}

Nas alíneas *b* e *c*, observamos o uso do “muito”, porém com características particulares de uso do PST. Na língua portuguesa, os pronomes indefinidos adjetivos são posicionados, tipicamente, antes do nome ao qual se referem. Nos exemplos em

²⁵ Ver contexto de uso na Figura 19 (p. 108).

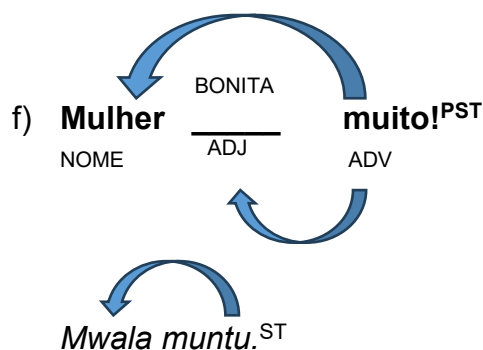
questão, no entanto, sucedem os nomes, posição comum do “muito” na língua Santome – *Naí sa ku ladlon muntu*.ST / *Ladlon muntu sa Santome*^{26ST}, respectivamente. Além disso, apresentamos o exemplo da alínea 11d – *San se tê wêxa muntu*.ST – e salientamos que a posição do *muntu* [muito] no ST é fixa e sempre após o nome, adjetivo ou advérbio com o qual se relaciona, independentemente da função desempenhada, seja de advérbio ou de pronome adjetivo.



O item lexical “dabuso” é resultado do processo de composição por aglutinação da preposição “de” e do nome “abuso”, mas manteve a sua etimologia – uso excessivo, demasiado – porém, no PST a conotação é neutra, ou seja, pode ser usado tanto para expressar coisas positivas como negativas – “Essa comida é boa dabuso.” ou “Aquele rapaz é feio dabuso”, respectivamente.

“Frontada”, por sua vez, é um adjetivo que foi emprestado do Santome – *flontadu* [flõ'tadu] (egoísta) e se aplica basicamente a todos os contextos em que alguém queira tudo ou a maior parte das coisas para si próprio. No exemplo sob análise, funciona como um advérbio de intensidade, pois está intensificando as características de alguém expressas pelo adjetivo “frontada”. Contudo, “dabuso” também pode ser usado como pronome indefinido para expressar quantidade – “Ladrão dabuso tá n São Tomé”.

²⁶ *Santome* nesse caso faz referência à Ilha de São Tomé ou ao país São Tomé e Príncipe.



No caso da expressão “Mulher muito!” (11f), embora o item lexical “muito” aparentemente desempenhe função de pronome adjetivo pelo fato de fazer referência a um substantivo (mulher), funciona, na verdade, como um advérbio de intensidade, pois está intensificando as características não expressas dela e é invariável, ou seja, existe um adjetivo com conotação positiva oculto entre o nome e o advérbio. Trata-se também de um empréstimo da língua Santome – *Mwala muntu!*, que é usado de forma irônica para expressar críticas ao interlocutor ou à pessoa a qual faz referência.

A expressão “Mulher muito!” foi usada em resposta à postagem “Nenhum homem de Trindade tem sabor.” A postagem faz uma crítica ao desempenho sexual dos nativos ou moradores da Cidade da Trindade (Distrito de Mé-Zóchi) na ilha de São Tomé. Nesse caso, “Mulher muito!” significa que a autora da postagem original não é tão bonita, por exemplo, quanto pensa. Aqui, a apresentação do contexto de uso e explicação se fazem necessárias para a compreensão geral da expressão uma vez que, na prática, os falantes se utilizam, paralelamente, de alguma expressão corporal e ideofone de desprezo.

Como podemos observar, com exceção da alínea 11a, em todos os demais casos apresentados no exemplo 11 (pp. 77-80), verificamos alterações das propriedades por

meio do contato linguístico do PST com ST. Portanto, o PST recebe influências do ST e, assim, compartilha de inúmeras das suas características lexicais e gramaticais.

A gramaticalização, contudo, pode se enquadrar em diversas abordagens. Para Gonçalves et al. (2007, p. 16), a gramaticalização pode ser um processo, paradigma, fenômeno diacrônico e sincrônico. Assim, a gramaticalização é considerada *paradigma* se observada num estudo da língua que se preocupe em focalizar a maneira como formas gramaticais e construções surgem e como são usadas. É considerada *processo* se se concentrar na identificação e análise de itens que se tornam mais gramaticais.

Neves (2021, p. 170) apresenta outras definições de gramaticalização segundo diferentes estudiosos e afirma que não há consenso. Ressalta ainda o uso de uma variante – gramaticização, por alguns autores – numa coletânea de Traugott e Heine, mas esclarece que Hopper e Traugott (1993) fazem uma distinção conceitual em que observam o uso de gramaticalização mais para uma “perspectiva histórica” e gramaticização, com uma “visão sincrônica da mudança de categorias e significados”.

De todo modo, citando Traugott e Heine (1991), Neves (2021 p. 170) afirma que gramaticalização “se refere à parte da teoria da linguagem que tem por objeto a interdependência entre *langue* e *parole*, entre o categorial e o menos categorial, entre o fixo e o menos fixo na língua.” Os autores também apontam o uso inadequado de termos alternativos como “sintatização, dessemantização, descoramento/ descoloramento/ enfraquecimento/ desbotamento/ desvanecimento semântico, reanálise, condensação, redução etc.” no lugar de gramaticalização.

Heine et al. (1991b) afirmam que:

Até 1970, a gramaticalização era vista principalmente como parte da linguística diacrônica, como um meio de analisar a evolução²⁷ linguística, de reconstruir a história de uma determinada língua ou grupo de línguas, ou de relacionar as estruturas linguísticas do momento com os padrões anteriores do uso linguístico. [...] Um dos principais méritos dos estudos sobre gramaticalização depois de 1970 foi a atenção dada ao potencial que eles oferecem como um parâmetro explanatório para a compreensão da gramática sincrônica. A insatisfação com os modelos de descrição gramatical existentes à época forneceu um grande incentivo para buscar na gramaticalização como meio de ultrapassar as abordagens "estáticas" de análise gramatical, em particular o estruturalismo e a gramática gerativo-transformacional. (Heine et al., 1991b, pp. 10-11) (tradução nossa)

A gramaticalização também é vista como um processo unidirecional, ou seja, não pode ser revertida. Segundo Neves (2021, p. 173), essa unidirecionalidade da gramaticalização é tida geralmente como característica básica do processo, partindo-se do princípio de que uma mudança que se dá na direção específica de uma categorização “gramatical” mais elevada não pode ser revertida. Heine et al. (1991b, p. 4), por exemplo, afirmam que essa unidirecionalidade é uma propriedade intrínseca da gramaticalização – de uma unidade "menos gramatical" para "mais gramatical", e não vice-versa. Lembram também que alguns autores tentaram provar o contrário (Kahr 1976; Jeffers e Zwicky 1980; Campbell etc.) com foco em desgramaticalização entendido como a direção da gramaticalização é revertida, isto é, quando uma unidade mais gramatical se transforma em uma menos gramatical, ou regramaticalização que é quando formas sem quaisquer funções adquirem função gramatical.

Porém, Lehmann (1982, *apud* Heine et al. 1991b, pp. 4-5), argumenta, que ocorrências de desgramaticalização e regramaticalização seriam estatisticamente

²⁷ Heine et al. ressaltam que alguns termos utilizados podem causar confusão por serem empregados com significados distintos na Linguística e em outras áreas acadêmicas. Por exemplo, o termo "evolução," é usado aqui para referir-se a mudanças no desenvolvimento de unidades ou estruturas linguísticas de acordo com as suas tendências inerentes (cf. Svorou 1988: 213). Reforçam a importância de ter em mente que tratam aqui não da "evolução das línguas", mas sim de evoluções relacionadas a partes específicas das línguas. Como ou até que ponto essas evoluções afetam a estrutura geral das línguas não é o foco da presente tese em construção.

insignificantes e, em muitos casos, a literatura encontrada revela ser resultado do uso de uma análise inadequada. Portanto, a gramaticalização se dá numa única direção – de uma unidade "menos gramatical" para uma unidade "mais gramatical" ainda. No caso da regramaticalização, apresentamos o uso especial dos ideofones “ê” e “ô”, originalmente do ST e incorporados ao PST, mas cujos limites das fronteiras entre a lexicalização, gramaticalização e mudança de código nem sempre são claros.

A interjeição “tchau” [tʃ'aw] no PST, geralmente, pronunciada sem o som do T – [ʃ'aw], vem acompanhada do ideofone “ê” [e] e a resposta do interlocutor tipicamente se dá apenas pela repetição do ideofone “ê” (ver exemplo 12). Por sua vez, o ideofone “ô” é utilizado após o vocativo para chamar alguém e a resposta também consiste apenas na repetição do mesmo ideofone (ver exemplo 13). Em outras palavras, esses ideofones tipicamente não têm uma função gramatical a não ser atribuir uma pequena ênfase. Porém, nos casos apresentados, desempenham a função gramatical.

(12) Regramaticalização do ideofone “ê”

a) Interlocutor A: (T)chau ê!^{PST/ST} / (T)chau!^{PT}

b) Interlocutor B: Ê!^{PST/ST} (resposta) / (T)chau!^{PT}

(13) Regramaticalização do ideofone “ô”

a) Interlocutor A: Fidel ô!^{PST/ST} / Fidel!^{PT}

b) Interlocutor B: Ô!^{PST/ST} (resposta) / Oi!^{PT}

Citando Hopper (1991, pp. 19-20), Neves (2021, p. 174) afirma que definir gramaticalização se torna mais problemática se o objetivo for a investigação de uma

única língua, vista isoladamente, e que, de uma perspectiva de estudo de mais de uma língua, é possível chegar-se a regularidades emergentes que têm potencial para ser instâncias de gramaticalização. E acrescenta:

Certos tipos de itens lexicais evoluem, tipicamente, para clíticos gramaticalizados e para afixos. Essas mudanças são muito ricamente documentadas, e alguns exemplos são os casos de verbos de cópula e de verbos de movimento que se tornam morfemas aspectuais, bem como casos de nomes locativos que se tornam, eventualmente, afixos casuais. A lista de mudanças desse tipo em línguas africanas, oferecida por Heine e Reh (1984, pp. 269-281), é de relevância universal. (Neves, 2021, p. 175)

2.3. CODE-SWITCHING

Como já foi explicitado no Item 1.3 Línguas (p. 48), São Tomé e Príncipe tem hoje, oficialmente, três línguas nativas - Santome, Ngola e Lung'le. Além dessas, o Português como língua oficial e o Kabuverdianu, como língua não-nativa, transplantada pelos trabalhadores contratados de Cabo Verde durante a colonização. Portanto, um ambiente multilíngue.

Neste subitem, discutiremos *code-switching* em todas as formas, tipos, níveis e diferentes abordagens, segundo teóricos. No entanto, como veremos mais adiante, esse fenômeno será abordado à luz dos estudos relacionados ao contato de línguas e bilinguismo e plurilinguismo, uma vez que é nessa perspectiva que São Tomé e Príncipe se enquadra.

Até 1970, o *code-switching* era geralmente visto como interferência de uma língua na outra e considerado um desempenho bilíngue (ou multilíngue) imperfeito ou incapacidade de o falante utilizar apenas umas das línguas e era considerado tema periférico. Depois, *code-switching* ganhou status central e estudos exclusivos, sendo

considerado hoje como performance dos falantes bilíngues e políglotas. *Code-switching* é um fenômeno inerente à oralidade dos falantes bilíngues ou políglotas, que se dá geralmente em ambientes e situações informais e de descontração, podendo dar-se também em situações formais.

No âmbito linguístico, o fenômeno de *code-switching* - ou alternância de línguas – é definido por Grosjean (1982, pp. 145-146) como “o uso alternado de dois ou mais códigos por indivíduos bilíngues numa mesma interação conversacional”. Por sua vez, Nílep (2006, p. 6) afirma que Gumperz foi o mais influente nos estudos sobre *code-switching*. Segundo o autor,

Talvez nenhum linguista sociocultural tenha sido mais influente nos estudos da mudança de código que John J. Gumperz. Seu trabalho sobre alternância de código e contextualização tem sido influente nas áreas da sociolinguística, linguística antropológica e sociologia da linguagem. Grande parte do trabalho inicial de Gumperz foi realizado no norte da Índia (Gumperz 1958, 1961, 1964a, 1964b), centrado no Hindi e nos seus múltiplos dialetos. Gumperz (1958) descreve três níveis – dialetos de aldeia, dialetos regionais e hindi padrão – cada um dos quais pode ser composto por numerosas variedades e que servem funções diferentes. (Nílep, 2006, p. 6) (tradução nossa)

Ainda de acordo com Nílep (2006), Blom & Gumperz (1972) dividiram *code-switching* entre metafórico e situacional. Para os autores o metafórico diz respeito à mudança relacionada ao efeito comunicativo de discurso, isto é, às intenções do falante e aos sentidos por ele pretendidos. Por outro lado, o fenômeno era considerado situacional quando tinha a língua e o contexto social de uso como foco, onde cada código tem um papel específico no repertório discursivo do falante em função da situação ou dos interlocutores.

Por sua vez, o fenômeno é classificado por Dabène & Moore (1995) em intrasentencial (quando ocorre o uso de uma palavra ou expressão dentro da mesma

sentença), intersentencial (alternância de sentenças em línguas diferentes) e entre enunciados – alternar para outra língua após um longo período de uso da primeira língua, afirma Mozzillo (2009).

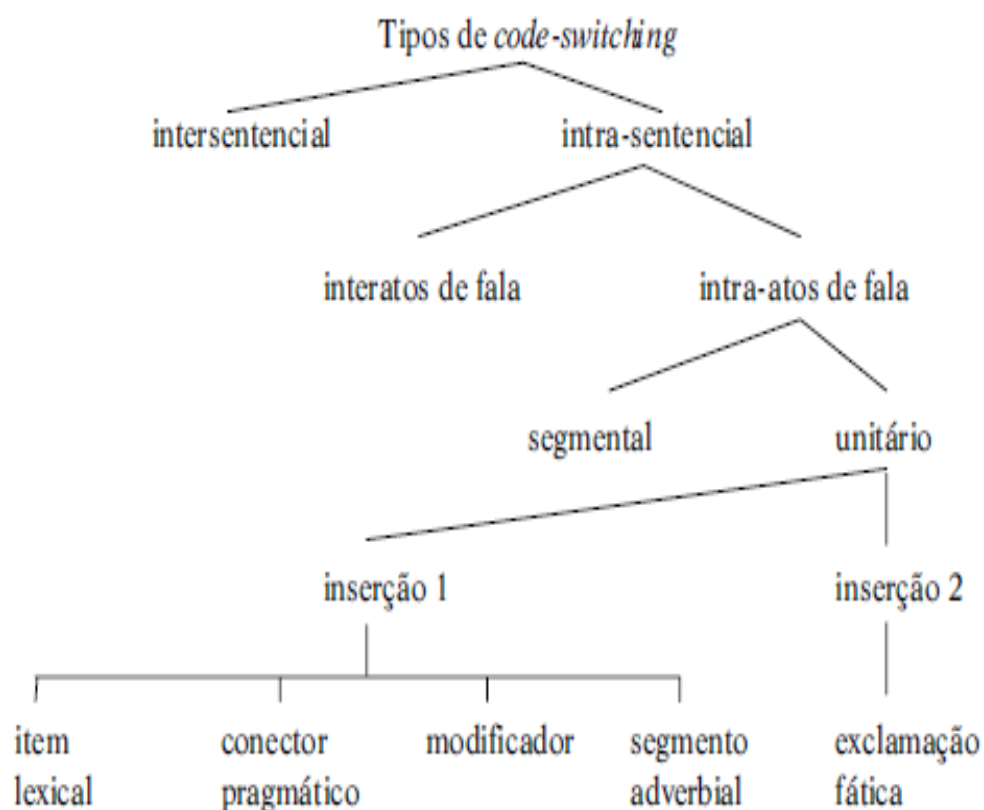
A mudança de código pode ser intencional ou involuntária. Os falantes fazem a mudança deliberada quando sentem necessidade, por exemplo, de preencher lacunas causadas pela falta de domínio completo da língua utilizada no momento ou quando querem estabelecer uma relação de proximidade/conexão com o seu público bilíngue. No caso de São Tomé (e Príncipe), isso é muito comum no período eleitoral. Durante as campanhas, quando fazem os seus comícios em regiões pesqueiras ou rurais, os candidatos utilizam-se do *code-switching* para demonstrar proximidade e pertencimento – um ato de identidade, para cativar eleitores.

Segundo Myers-Scotton (1993),

O falante emprega a noção de marcado e não-marcado para julgar quais códigos deve usar na conversação. Normalmente, o que for menos marcado será o que empregará, embora não o faça sempre porque pode, justamente, querer implicar coisas diferentes. Na maioria das vezes, suas escolhas e decisões quanto a que língua falar na relação interpessoal, são inconscientes. Segundo a autora, os falantes sabem, como parte de sua competência comunicativa, que escolher um código implica negociar identidades na interação em pauta e, por isso, levam em conta as consequências. (Myers-Scotton *apud* Mozzillo, 2009, p. 188)

Mozzillo (2009) sustenta também que entre as várias classificações de *code-switching* uma das mais abrangentes é a de Dabène & Moore (1995). Segundo as autoras, *code-switching* pode ser intra-sentencial, intersentencial e entre enunciados (ver Figura 14, p. 87).

Figura 14: Tipos de *code-switching*



Fonte: Mozzillo (2009, p. 188)

Na Figura 15 (p. 88), apresentamos um exemplo de alternância de código intersentencial, ou seja, o usuário só muda de PST para ST depois de concluir a sua frase no primeiro idioma. Aqui o recurso ao *code-switching* ocorre porque se trata de uma questão de ditado popular (os “versos”), muito utilizados em São Tomé e, tipicamente no Santome. Essa alternância se faz necessária por necessidade de expressão do pensamento que só fica completa, sarcástica e ao mesmo tempo irônica no ST:

(14) *Tason kume inen jêlu, bil[a]*²⁸ *paga inen ku mali.*ST

Sentar comer 3PL dinheiro voltar pagar 3PL PREP
mal

Desfrutaram do dinheiro deles e ainda fizeram mal a eles.^{PT}

Figura 15: *Code-switching* intersentencial



Fonte: Facebook (2023)

²⁸ Embora a palavra *bili* também exista no Santome, ela tem sentido de “abrir”. Por sua vez, “*bila*” significa “voltar”, “virar”, “girar” e é usada também como locução adverbial “de novo”, “novamente” etc. Portanto, por se tratar de erro de ortografia, realizamos a correção necessária.

2.4.A TESE

O título desta tese não é aleatório; foi cuidadosamente pensado – **“Português de São Tomé (e Príncipe): lexicalização, gramaticalização e *code-switching* em contexto multilíngue”**, porque representa um posicionamento político do autor, à medida que até para a maioria dos são-tomenses (letrados, elite) trata-se apenas de um “falar” são-tomense, um português mal falado.

A hipótese que sustenta a tese decorre da observação de que as línguas faladas em São Tomé e Príncipe resultam do contato entre o português de Portugal e as línguas originais do STP. Em vista disso, procederemos à análise de dados que demonstrem que 1) processos de lexicalização e 2) processos de gramaticalização do português de São Tomé (e Príncipe) resultam de empréstimos de itens lexicais e de decalques de estruturas sintáticas do Santome com vistas a demonstrar que a compreensão do português falado em São Tomé (e Príncipe) depende, muitas vezes, do conhecimento do Santome na base da primeira língua. Portanto, para alcançar os objetivos propostos nesta tese e verificar a hipótese levantada, apresentaremos em seguida, no Capítulo 3, os recursos, estratégias e os procedimentos metodológicos adotados.

CAPÍTULO 3: METODOLOGIA DE PESQUISA

3.1. PROCEDIMENTOS METODOLÓGICOS

Neste capítulo, apresentamos os procedimentos metodológicos da pesquisa. Antes disso, porém, importa tecer algumas observações sobre a desambiguação dos itens lexicais “São Tomé” e “Santome” [sɛ̃to'mɛ] em razão da homofonia.

A palavra São Tomé é polissêmica e mantém uma relação de sinonímia com Santome, Forro e dialeto. Dependendo do contexto de uso, São Tomé poderá designar a Ilha de São Tomé, o país São Tomé e Príncipe ou ainda Santome, crioulo Forro, também chamado de “dialeto” pelos são-tomenses. Além disso, no próprio Santome existem variantes como *Fôlô* [Forro], *Lungwa Santome* [Língua de São Tomé] e *Lungwa Tela* [Língua da Terra].

Cabe considerar, na continuidade dessa discussão, o que é variante. Para TARALLO (1986, p. 08) “variantes linguísticas são diversas maneiras de se dizer a mesma coisa em um mesmo contexto e com o mesmo valor de verdade. A um conjunto de variantes dá-se o nome de *variável linguística*”. Assim, explicar a desambiguação de São Tomé/Santome se faz necessária por três motivos:

- 2) Santome é termo oficialmente fixado em São Tomé e Príncipe pelo Decreto de nº 19, de 14 de agosto de 2013, que dispôs sobre a ortografia e criou o Alfabeto Unificado para as Línguas Nativas de São Tomé e Príncipe (ALUSTP). Portanto, é ele que é utilizado nesta tese.
- 3) Na prática, não existe diferença fonética na pronúncia de “São Tomé” (ilha, capital ou país) e Santome (língua, ilha, capital ou país).

- 4) Existem diversos trabalhos acadêmicos sobre Santome – alguns trazem as duas designações (Santome e Forro, na mesma publicação), outros utilizam ou Forro ou Santome, mas raramente encontramos algum trabalho que faça essa distinção de variação conceitual referindo-se ao Forro como grupo étnico, e não apenas como uma língua crioula.

3.2. A PESQUISA: TIPO E NATUREZA

Os procedimentos metodológicos adotados nesta pesquisa são de natureza qualitativa, uma vez que busca descrever o contexto sócio-histórico-cultural-linguístico do processo de formação da sociedade são-tomense, para em seguida, analisar os processos de mudança linguística do Português de São Tomé (PST) com foco na lexicalização, gramaticalização e no *code-switching*.

Em segunda ordem, a pesquisa adota uma abordagem descritiva, já que apresenta a descrição dos processos e fenômenos linguísticos no PST. Portanto, para alcançar os objetivos propostos, os seguintes procedimentos foram adotados: 1) revisão da bibliografia; 2) coleta de dados e formação de corpora; 3) tratamento e análise dos dados à luz das teorias pertinentes e conclusões em face dos resultados das análises.

3.3. LEVANTAMENTO BIBLIOGRÁFICO

Depois de delimitado o escopo e limites da pesquisa, foi feito um mapeamento geral do estado da arte sobre os trabalhos que tratavam do tema, dos mais antigos até

aqueles que mostram as tendências atuais sobre lexicalização, gramaticalização e *code-switching*.

Para a fundamentação teórica, seguiremos uma abordagem funcionalista de mudança linguística, pois consideramos os aspectos dinâmicos da língua envolvendo os processos cognitivos e fatores inerentes ao uso da língua em situações comunicativas pelos falantes. Para tanto, nos apoiaremos, principalmente, nos(as) seguintes autores(as): Bybee - “Língua, uso e cognição” (2016 [2010]) e “Mudança Linguística” (2020 [2015]); Moura Neves - “Texto e Gramática” (2016) e “Gramática Funcional: interação, discurso e texto” (2021).

Consideramos aqui que, no processo de mudança linguística, a mudança lexical precede a mudança gramatical ou ajuste na função gramatical. O *code-switching*, por sua vez, também se enquadra antes de tudo no nível lexical e, no caso de São Tomé (e Príncipe), o Português é a L1 para enorme maioria da população e a alternância de código acontece principalmente na sequência Português>Santome, ou seja, enquanto se fala português, recorre-se ao Santome. No quesito de centralidade do léxico, Faulstich (2012, p. 368) argumenta que:

Toda língua possui um fundo lexical, que constitui um dos fundamentos da expressão linguística e que representa o “dicionário mental” do indivíduo. O fundo lexical é, por conseguinte, um componente no qual se acumulam todos os elementos léxicos de uma língua – predicados e palavras –, assim como as regras, por meio das quais é possível criar novas entidades de um modo produtivo. (Faulstich, 2012, p. 368)

O corpus é predominantemente do domínio da oralidade. Assim sendo, os corpora constituídos são registros da oralidade. Inicialmente com previsão de coleta dos dados *in loco*, a pesquisa voltou-se para os registros (diálogos e afins) em livros publicados e principalmente nas postagens nas mídias sociais Facebook e YouTube. Com o avanço

da tecnologia e normalização do universo digital, as mídias digitais se transformaram em plataformas de interação social e são representativas de comunidades de fala. Para fundamentar este conceito, consideramos o que diz Labov (1972):

A comunidade de fala não é definida por nenhum acordo/contrato no uso de elementos de língua, mas pela participação em um jogo de normas compartilhadas; tais normas podem ser observadas em tipos claros de comportamentos avaliativos e pela uniformidade de modelos abstratos dos padrões da variação que são invariáveis em relação aos níveis particulares de uso. (Labov, 1972, pp. 120-121)

Esse conceito de comunidade de fala de Labov é bastante pertinente porque os dados coletados mostram que exemplos de lexicalização e de gramaticalização assim como alternância de código no uso da língua nas comunidades são-tomenses seguem uma uniformidade abstrata de padrões da variação invariáveis nos usos individuais, isto é, normas compartilhadas, conforme a Figura 16 (p. 94):

Figura 16: Postagem de CS



Fonte: Elaborado pelo autor

(15) **Gente, eu salguei gui[s]ado de minha mãe. Como é que eu faço para reduzir o sal?**

- a) Comentários 1 e 11: Coloca água
- b) Comentários 2 e 8: Aumenta água
- c) Comentários 5, 8 e 12: Larga água
- d) Comentários 6 e 7: Põe água
- e) Comentário 10: Bota água
- f) Comentário 11: Ferve água
- g) Comentário 13*: Folha quime

Nos exemplos acima extraídos da Figura 16 (p. 94), identificamos variação nas opções de verbos possíveis para reger o nome “água”, limitada no contexto de uso apresentado. Para efeito de análise, deixaremos de fora os constituintes não essenciais dos comentários 1, 7, 8 e 12. Assim sendo, teremos o seguinte: colocar + água, aumentar + água, largar + água, pôr + água, botar + água e *ferver + água. Em português padrão, todas as combinações parecem naturais exceto, “largar água”. Desse modo, os comentários 5, 8 e 12 seriam, provavelmente, incompreensíveis fora do contexto de uso, ou até mesmo no contexto, para alguém que não seja são-tomense ou more em São Tomé, a comunidade de fala são-tomense. Esses casos são, na verdade, empréstimos feitos pelo português de São Tomé (PST) por meio de decalque do Santome (ST) – *lege awa*, cujo sentido é “aumentar água”. No caso do comentário 8, temos um ideofone “*eeee*” /εεεε/ que também é um empréstimo do ST *êêêê* /eeee/, cuja função é

basicamente enfatizar a ação do verbo – *Lege awa éêêê*. A alínea *e* (botar água), por sua vez, também não é original de STP, pois é um empréstimo do Português do Brasil.

Além dos pontos abordados no parágrafo anterior, apresentamos também extrato do comentário 13, “folha quime”. Quando se trata de designar folhas das plantas, em português, a regra é a utilização da preposição “de”, cuja função é estabelecer a ligação entre duas palavras ou termos: folha de bananeira, folha de goiabeira, folha de jaqueira etc. No PST, no entanto, o mais comum é a ausência dessa preposição pela contribuição e impacto do ST: folha bananeira (*fya-bana*ST), folha fruteira (*fya-fluta*ST), folha jaqueira (*fya-jaka*ST), folha kimi (*fya-kimi*ST, bot. *Newbouldia laevis*), folha micocó (*fya-mikoko*ST, Alfavaca-cravo ^{PT}), folha ponto (*fya-pontu*ST, *Achyranthes áspera*), *fya-gleza*²⁹ (samambaia) etc. Por fim, no comentário 4, temos o *Ke kwa!* (grafia segundo ALUSTP), uma interjeição emprestada do ST, que é usada para expressar espanto, geralmente com conotação negativa.

Na explicitação da Figura 16, observamos ocorrências de lexicalização (léxico), com empréstimos de palavras do ST, gramaticalização (estrutura sintática), novo verbo (largar) que, embora existente previamente no português, expandiu a sua aplicação e regência e incluiu o sentido de “colocar, por e aumentar” e a queda da preposição “de” no estabelecimento de relação entre dois nomes além da ocorrência de *code-switching*, por meio do uso da interjeição *Ke kwa!*, do ST. Em suma, verificamos a complexidade cognitiva e linguística nas comunidades de fala são-tomenses e o impacto do ST no PST.

²⁹ *Fya-glêza* (bot. *Dryopteris filixmas*. Samambaia) tem o sentido literal de “folha-de-igreja”, pois é usada na construção do “Paço Fya-Glêza”, em prevalece a designação “fya-glêza”, do Santome. “Paço Fiá Gleza, é o presépio de Natal em São Tomé e Príncipe. É a principal referência cultural são-tomense na celebração do dia da família, o dia em que a maioria absoluta do povo são-tomense celebra o nascimento de Cristo. Na berma da estrada são construídos os Paços Fiá Gleza que embelezam os bairros, sobretudo no interior do país.” (Tela Non, 2010)

Portanto, dada a amplitude das características peculiares de São Tomé (e Príncipe), alargaremos a discussão na Linguística Centrada no Uso (*Usage-based Linguistics*) para expandir os procedimentos e estabelecer uma interface com *code-switching*. A situação linguística de São Tomé é tão complexa, e tudo tão ‘misturado’, que é quase impossível delimitar fronteiras linguísticas e sociais e, conseqüentemente, separar ou restringir claramente o arcabouço teórico-metodológico para o trabalho aqui proposto.

3.4. DADOS: DELIMITAÇÃO E COLETA

Para Dayrell (2012, p. 87), corpus é “uma coleção de textos selecionados e agrupados de acordo com critérios claramente definidos e especificados.” Nesse sentido, nesta tese, a coleta de textos se organizou assim:

- a) Do Facebook foram coletadas postagens principalmente em seis (6) grupos relacionados a São Tomé e Príncipe (nomeadamente Acorda Santomenses, Espaço Santomense, Voz de S.Tomé e Príncipe, Pro-São-tomenses, Santomemes), Kua Non e os respectivos comentários assim como postagens aleatórias feitas por usuários são-tomenses e os comentários feitos por outros(as) são-tomenses.
- b) Textos extraídos dos seguintes livros: Histórias da Roça (Fernando Reis, 1970), Contos Tradicionais Santomenses (Dire(c)ção Naciona da Cultura de São Tomé e Príncipe, 1984), A Ilha de São Tomé nos Séculos XV e XVI (Luis de Albuquerque, 1989), Semplu

(Olinto Daio, 2002), Equador (Miguel S. Tavares, 2004), Sóias e Contági (Teresa Macara & Luísa Lisboa, 2009), Soyas de Sun Tataluga (Rogério A. Barbosa & Taísa Borges, 2015).

- c) Matérias publicadas em alguns jornais digitais locais: Tela Non, STP Press, PNUD São Tomé e Príncipe e VOA Mais África e o Mundo.
- d) Do YouTube, vídeos de programas da TVS (Televisão São-tomense) como “Nós por cá”³⁰ assim como videocliques das músicas são-tomenses que apresentam ocorrências de dados importantes para esta pesquisa.

Importa ressaltar ainda que grande parte das letras de músicas de STP produzidas no Português de São Tomé (e Príncipe) não está disponível na internet e, tampouco, as letras de músicas produzidas no Santome estão disponíveis em algum lugar. Portanto, a pesquisa demandou mais tempo e trabalho, pois exigiu transcrição das letras dessas músicas uma vez que nesses vídeos estão presentes muitas ocorrências, principalmente de *code-switching*, assim como itens lexicais e estruturas sintáticas que constam atualmente como empréstimos ou neologismos no PST.

³⁰ Peças de Teatro feitas e apresentadas em São Tomé, pelo grupo tradicional são-tomense Os Criativos. Do outro lado, existe também o “Nós por lá”, entrevistas feitas por são-tomenses a outras pessoas naturais ou de origem são-tomense que vivem na diáspora.

3.5. SÍNTESE DO CAPÍTULO

Em suma, esta pesquisa é de natureza qualitativa e de abordagem descritiva; propõe-se a apresentar o contexto sócio-histórico e linguístico de São Tomé e Príncipe (STP) e a descrever os fenômenos linguísticos e processos de mudança linguística no arquipélago, por meio de dados coletados principalmente nas mídias sociais. Assim sendo, concentramos especificamente nas particularidades do Português de São Tomé (PST), português falado na ilha de São Tomé, e o contrastamos com a língua nativa Santome para descrever os processos de formação de itens lexicais (lexicalização), criação de estruturas sintáticas totalmente novas e/ou mudança de estruturas existentes (gramaticalização), seja resultante do novo item lexical seja mudança de categoria de um item já existente, e *code-switching* ou uso alternado das línguas português-santome, por falantes nativos. Para tanto, nos baseamos nas abordagens funcionalistas de mudança linguística e nos pressupostos linguísticos de línguas em contato, uma vez que o PST está inserido em um contexto multilíngue.

CAPÍTULO 4: ANÁLISE DOS DADOS

4.1. PROCEDIMENTOS DE ANÁLISE

Neste capítulo, apresentamos os dados analisados à luz das teorias apresentadas. As análises seguem a sequência metodológica léxico>gramática>*code-switching* para delimitar e categorizar os exemplos em léxico, gramática e *code-switching*, embora muitos itens lexicais se enquadrem em mais de uma categoria. Nesses casos, todos os processos identificados nesses itens lexicais serão analisados na categoria inicial, mas poderão ser (re)analisados na(s) nova(s) categoria(s), observados os critérios específicos de cada categoria de análise – lexicalização, gramaticalização e *code-switching* caso seja pertinente.

Além disso, serão utilizados alguns símbolos matemáticos, listados abaixo, com o objetivo de facilitar ou simplificar a análise dos itens lexicais:

- a) + para indicar soma ou junção de partes constituintes de uma palavra ou item lexical;
- b) = para indicar o resultado da soma ou junção de partes constituintes de uma palavra ou item lexical; o símbolo de igualdade será utilizado também para indicar sinonímia na mesma língua.
- c) $\equiv$ (equivalente a) para indicar tradução ou item lexical correspondente no português (PT).
- d) PT será utilizado quando as palavras ou itens lexicais forem os mesmos tanto no PST como no Português do Brasil (PB).

- e) $\cong$ (aproximadamente igual) para indicar semelhança entre os significados de duas palavras ou itens lexicais no ST e PST, ou quando não houver um correspondente direto da palavra ou item lexical do ST no PST, PT.
- f) $>$ para indicar a transferência/transição, evolução, mudança de uma palavra, item lexical ou estrutura sintática do Santome (ST) para o Português de São Tomé (PST), ou apenas dentro do PST, com ênfase no processo.
- g) $<$ para indicar a transferência/transição, evolução, mudança de uma palavra, item lexical ou estrutura sintática do Santome (ST) para o Português de São Tomé (PST), com ênfase no resultado.
- h) PB será usado para designar palavras, termos ou expressões correspondentes no Português do Brasil quando tais palavras, termos e expressões no PST forem diferentes do PB.
- i) Para análise dos dados das postagens, serão utilizadas apenas as iniciais dos nomes dos(as) usuários(as) da mídia responsáveis pelas postagens e as imagens dos perfis serão desfocadas para proteger as identidades.

A tentativa de apresentação prévia de uma lista completa dos itens lexicais selecionados e das questões a serem analisadas mostrou-se impraticável e ineficiente, pois teríamos que antever todos os fenômenos e processos linguísticos. Assim sendo, optamos por agrupar e analisar os itens lexicais que apresentavam características semelhantes. Portanto, após a explanação dos procedimentos, passamos à apresentação, descrição e análises dos exemplos extraídos dos dados coletados.

4.2. LEXICALIZAÇÃO

4.2.1. Análise de itens lexicais sob o ponto de vista da morfologia

Como afirmamos anteriormente, a sociedade são-tomense é multilíngue. Porém, o ST, como uma das duas línguas predominantes além do Português, se faz presente basicamente em todas as áreas do cotidiano dos são-tomenses – fauna, flora, culinária, política, turismo, medicina, toponímia etc. Nesse sentido, apresentamos, primeiramente, alguns itens lexicais do PST que foram emprestados do ST ou por não terem designação própria no PST, ou porque são de uso mais frequentes ou simplesmente por concorrerem com seus correspondentes no PST.

4.2.2. Nomes comuns emprestados do Santome sem alteração

- (16) *Xikila*^{ST/PST} [ʃiki'la] (n.) Cooperativa em que, a cada mês, algum membro dá uma parte dos seus vencimentos para ser guardado num fundo que poderá ser utilizado futuramente.

Grafia proposta para o PST → “chiquilá”

- (17) *Kota-bega*^{ST/PST} ['kɔta 'bɛga] / *Kode*^{ST/PST}(var.) (n.) = Caçula

Grafia proposta para o PST → “cotabega” / “Codé”

- (18) *Ôkôsô*^{ST/PST} ['okoso] (n.) = Albino

Grafia proposta para o PST → “ôcosso”

- (19) *Gugu*^{ST/PST} [gu'gu] (n.) = Duende. (Figura mitológica)

Grafia proposta para o PST → “gugu”

- (20) (Casa de) *Vamplega* [vẽple'ga] ou *vaplega* [veple'ga] – habitação modesta construída com *andala*^{ST/PST} (folha de palmeira) nas roças, sendo que a *banca*^{ST/PST} (pecíolo^{PB}) é utilizada para fazer as paredes e a *andala*³¹, trançada, é usada para cobrir a casa.
- (21) *Sôtxi-flima* ['sotʃi fli'ma] (n.) = *Sôtxi-flima*. Ritual antigo de transição da puberdade para a vida adulta.

Após a apresentação dos diversos itens lexicais emprestados do ST e inseridos no PST sem alterações gráficas, passamos à análise morfológica, centrada no léxico, dos itens lexicais a seguir. Adotamos o conceito de Nida (1970, p. 1), “o estudo dos morfemas e seus arranjos na formação das palavras” e os itens lexicais selecionados serão analisados segundo os arranjos dos morfemas na formação das palavras.

4.2.3. Itens lexicais resultantes de processos parciais de mudança

linguística

- (22) *Ope pau*
- (23) *Pau bôbô*³²
- (24) *Pé kwali*
- (25) *Barba mixidaji*
- (26) Calça *sat'awa*³³

³¹ a) *Andala* é utilizada no ST (*ndala*) e no PST para se referir tanto à folha de palmeira como à folha do coqueiro. Porém, a (casa de) *vamlegá* é feita geralmente com a *bança*^{PST} (pecíolo^{PB}) da *andala* da palmeira. b) A palmeira pequena/jovem é chamada de *mutendé* no ST e muitas vezes no PST também.

³² *Pau bôbô*^{PST} ou *pau bobodu*^{PST} pode ser literalmente traduzida como “pedúnculo maduro”.

³³ Em português, foram encontrados termos aproximados como “calça capri” e “calça pescador”. Porém, em São Tomé e Príncipe, *sat'awa* refere-se à calça social masculina, geralmente feita em tecido ou sarja, cujo comprimento fica um pouco abaixo do meio da canela, cerca de cinco centímetros (5cm) acima do tornozelo.

Os itens lexicais acima são todos substantivos compostos e, curiosa e interessantemente, formados por duas palavras em duas línguas – ST e PST, razão pela qual chamamos aqui de mudança parcial. *Ope pau*^{PST} [perna de pau^{PB}] é um item lexical que entrou no PST por empréstimo do ST, *ope po*ST, mas que foi parcialmente traduzido para PT, ou seja, o primeiro item *ope* [pé, perna^{PT}] permaneceu no Santome e apenas o segundo componente foi transferido para português *pó* [pau^{PT}]. De modo similar, *pau bôbô*^{34PST}, *pé kwali*^{PST} [pé cesto^{PST}, perna arqueada^{PB}] e *barba mixidagi*^{PST} [≅ barba falhada^{PT}] também foram emprestados do Santome, mas apenas a primeira parte desses itens lexicais foi aportuguesada e a segunda permaneceu no Santome - *po bôbô*ST ou *po bôbôdu*ST, *ope kwali*ST e *beba mixidagi*ST. Temos ainda a calça *sat'awa* ou simplesmente *sat'awa*³⁵ [≅ calça capri, pescador^{PB}]. Utilizam-se ambas as variantes escritas porque as duas são possíveis e por não existir uma ortografia formal e oficial para o item lexical. *Satawa* ou *sat'awa* deriva de duas palavras *satá*ST /sa`ta/ [= saltar, pular<sup>PT</sup>] e *awa*<sup>ST</sup> /`a.wa/ [água^{PT}]. Portanto, *sat'awa* significa literalmente “saltar a água” – andar na água sem molhar a pé da calça.

4.2.4. Itens lexicais resultantes de processos de mudança linguística por derivação

(27) Fulumento (Figura 17, p. 105)

³⁴ *Pau bôbô*^{PST} ou *pau bobodu*^{PST} pode ser literalmente traduzida como “pedúnculo maduro”.

³⁵ Em português, foram encontrados termos aproximados como “calça capri” e “calça pescador”. Porém, em São Tomé e Príncipe, *sat'awa* refere-se à calça social masculina, geralmente feita em tecido ou sarja, cujo comprimento fica um pouco abaixo do meio da canela, cerca de cinco centímetros (5cm) acima do tornozelo.

Figura 17: Fulumento



Fonte: Facebook (2025)³⁶

O item lexical *fulumento* é composto pelo morfema lexical (radical) *fulu* do ST e o morfema gramatical (sufixo) de formação de nomes de ação do PT *-mento*. Mas, sob a perspectiva etimológica, no ST, *fulu*³⁷ é um verbo e significa “meter a mão para escolher algo afobadamente”.

Na Figura 17, temos um ideofone “e” /ɛ/ que também é um empréstimo do ideofone ê /e/ do Santome, cuja função é basicamente enfatizar a ação do verbo oculto

³⁶

Disponível

em:

<https://www.facebook.com/photo.php?fbid=1278039844328905&set=pb.100063687037480.-2207520000&type=3>. Acesso em 29 de maio de 2025.

³⁷ *Fulu* também é usado como verbo no PST com o mesmo sentido de Santome. Não confundir *fulu* com *Fulu-Fulu*, um peixe da família do atum.

“fazer” – Nada de [fazer] fulumento é, freguesa.^{PST} *Na fe fulumento fa ê, flêgê.*ST Da mesma forma, verificamos o uso do mesmo ideofone na segunda construção – “Vem comprar fruta é, freguês.”^{PST} – *Bi kompla fluta ê, flêgê.*ST Portanto, o ideofone “e” /ɛ/ do PST é o resultado da gramaticalização do ideofone ê /e/ do Santome, e preservam a mesma função e posição tanto no eixo paradigmático como sintagmático.

- a) Nada de [fazer] fulumento é, freguesa.^{PST}
 Nada de [fazer] fulumento IDEO freguesa
*Na fe fulumento fa ê, flêgê. Bi kompla fluta ê, flêgê.*ST

➤ ≅ Não precisa brigar, freguesa.^{PT}

- b) Vem comprar fruta³⁸ é, freguês.”^{PST}
 VIR comprar fruta IDEO freguês
*Bi kompla fluta ê, flêgê.*ST

➤ Vem comprar fruta-pão, freguês.^{PT}

(28) Mafeismo

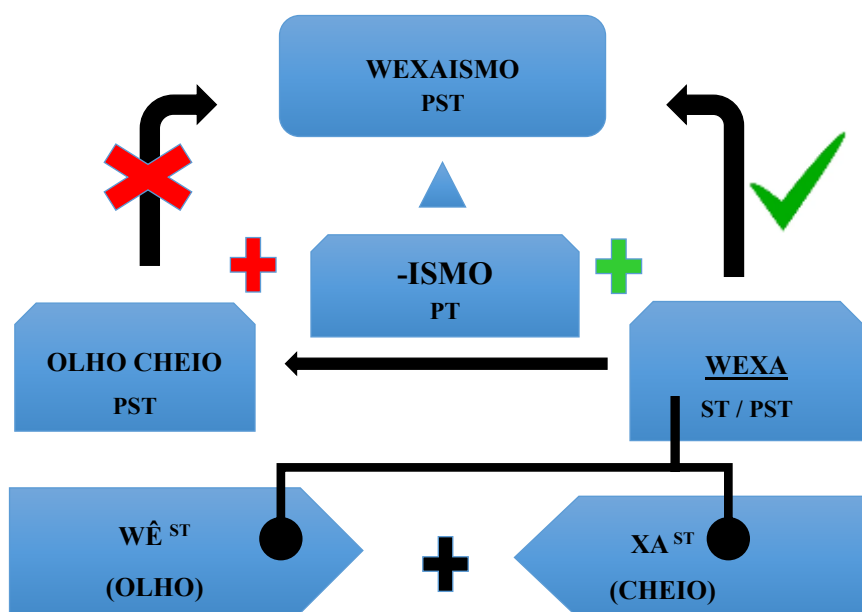
(29) wêxaismo

Os itens lexicais *mafeismo* e *wêxaismo* sofreram igualmente mudança com inserção do sufixo de formação de nomes “ismo”. Porém, *mafe* e *uêcha* são raízes de origem bem diversas. *Mafe* é um radical criado pela justaposição do substantivo comum

³⁸ Em São Tomé, o item lexical “fruta” é usado rotineiramente para se referir à fruta-pão. Do mesmo modo, “fruteira” é usada para fazer referência à árvore da fruta-pão (*Artocarpus altilis*).

português “má-fé”. Por sua vez, o radical *uêcha* é do Santome e é o resultado da justaposição das palavras *uê* e *cha* (olho cheio³⁹), que significa “inveja”. Além disso, foram identificadas ocorrências de variação ortográfica para esse item lexical – *uechaismo*, *uexaismo* e *wechaismo* – assim como para *uê cha* – *uê xa* e *wê xa*. E, embora item lexical *uechaismo* anteceda o ALUSTP, esteja consolidado e seja a forma mais usada pelos falantes, no ALUSTP, o fonema /j/ é representado pelo grafema “x”. Portanto, adotamos a grafia *wexaismo* (Figura 18) neste trabalho.

Figura 18: Wexaismo: processo de formação



Fonte: Elaborado pelo autor

A Figura 18 (p. 107), revela, na verdade, grande concorrência no processo cognitivo entre os itens lexicais “olho cheio” e “wê xa” no PST, pois ao invés de se

³⁹ Em São Tomé, as pessoas menos letradas e/ou menos afortunadas tendem a utilizar “olho cheio” no lugar de “inveja” e isso revela-se como um impacto direto do Santome no PST e criou um sinônimo concorrente para a palavra “inveja”.

lexicalizar por meio de “olho cheio”, o termo se derivou do item lexical “wêxa” do Santome. Porém, a pesquisa mostra uma variação maior e inclui outros sinônimos – olho grosso, olho gordo e inveja. Importa ressaltar também que, apesar dessa variação, o item “invejoso” e suas flexões ainda são praticamente os únicos usados na função de sujeitos para pessoas com tais características e na função de adjetivos para indicar atributos (ver Comentários 9 e 10 da Figura 19).

Figura 19: Olho cheio e suas variações

Comentário 1 A bolsa é mesmo de homem. Vai <u>mazé</u> procurar o que fazer.		Comentário 8 <u>Olho cheio</u> que te faz. Deixa o jovem em paz.
Comentário 2 Olho cheio começa assim.		Comentário 9 <u>Xé!</u> O que tem uma saca? [...] E vocês <u>invejosos</u> vindo cá falar mal de uma saca. [...] Vão <u>mazé</u> comprar a vossa saca. Isso é <u>inveja</u>.
Comentário 3 <u>Inveja</u> mata, mas os honestos vencem sempre.		Comentário 10 Santomense já começou com <u>olho gordo</u>...<u>invejosos</u>...
Comentário 4 A bolsa ficou muito bem. Mano, <u>inveja</u> mata.		Comentário 11 <i>Tudu kwa té wê!</i>
Comentário 5 Vai se entregar na igreja, <u>olho grosso</u>.		Comentário 12 Livro, vai estudar, <u>vajin</u>.
Comentário 6 <i>San se té <u>wê xa muntu</u>.</i>		Comentário 13 Homem e muito homem. Vai lavar prato com a tua <u>inveja</u>.
Comentário 7 <u>Olho cheio</u> <u>mais é</u>.		

Fonte: Elaborado pelo autor

Para explicar os processos de lexicalização e de gramaticalização de *fulumento* e *uechaismo*, seguimos Andrade et al. (2019) quando afirmam que a existência de contato intensivo entre línguas tende a resultar em aspectos linguísticos altamente variáveis. E Calvet (2002, p. 54) reforça:

Varição e mudança são fenômenos naturais nas línguas do mundo e que as línguas crioulas, por manifestá-los com tanta pujança, são importantes para a compreensão de fatos dessa natureza, bem como para a compreensão “da linguagem humana” como um todo. (Calvet, 2002 p. 54)

Outra questão importante é o item lexical “mazé” nos comentários 1, 7 e 9 (ver exemplos 30, 31 e 32, respectivamente), embora no comentário 9 tenha sido usada a forma original (“mais é”), o que pode ser visto como um processo de mudança linguística em curso no PST. “Mazé” é resultado do processo de composição por aglutinação do adverbio de intensidade “mais” e a terceira pessoal do singular do verbo ser “é”. Nesse processo, porém, ocorreu a sincopa do fonema “i” no interior da palavra seguido da aglutinação do “mas + e” (mais é > mas’e > mazé):

(30) Vai mazé procurar o que fazer.

≅ Vai procurar o que fazer, isso sim.^{PT}

(31) Olho cheio mais é.

≅ Inveja, isso sim.^{PT}

(32) Vão mazé comprar a vossa saca.

≅ Vão comprar a vossa bolsa, isso sim.^{PT}

Portanto, no PST, o item lexical “mazé”, derivado da expressão “mais é”, é usado pelo falante geralmente para contrastar com um evento de fala anterior do seu interlocutor conforme os exemplos 30, 31 e 32. Na argumentação sintática, a expressão

é posicionada tipicamente após o verbo ou locução verbal (30 e 32) e no final da frase (31), pois a sentença completa no 31 é “Isso é olho cheio, mais é.” Em outros contextos, a expressão pode ser usada para contrariar e com o sentido de “na verdade” – Vocês não gostam de estudar, maze. [Na verdade, vocês não gostam de estudar.]

Ainda na Figura 19 (p. 108), temos o nome do usuário autor da postagem – *Peixe Pambi*. Esse é um item lexical muito particular do PST, pois está vinculado à história de São Tomé e Príncipe, principalmente à pós-independência, popularmente conhecido como o período da Primeira República (1975-1991), que era caracterizado pelo partido único, desafios político-sociais e grande insegurança alimentar.

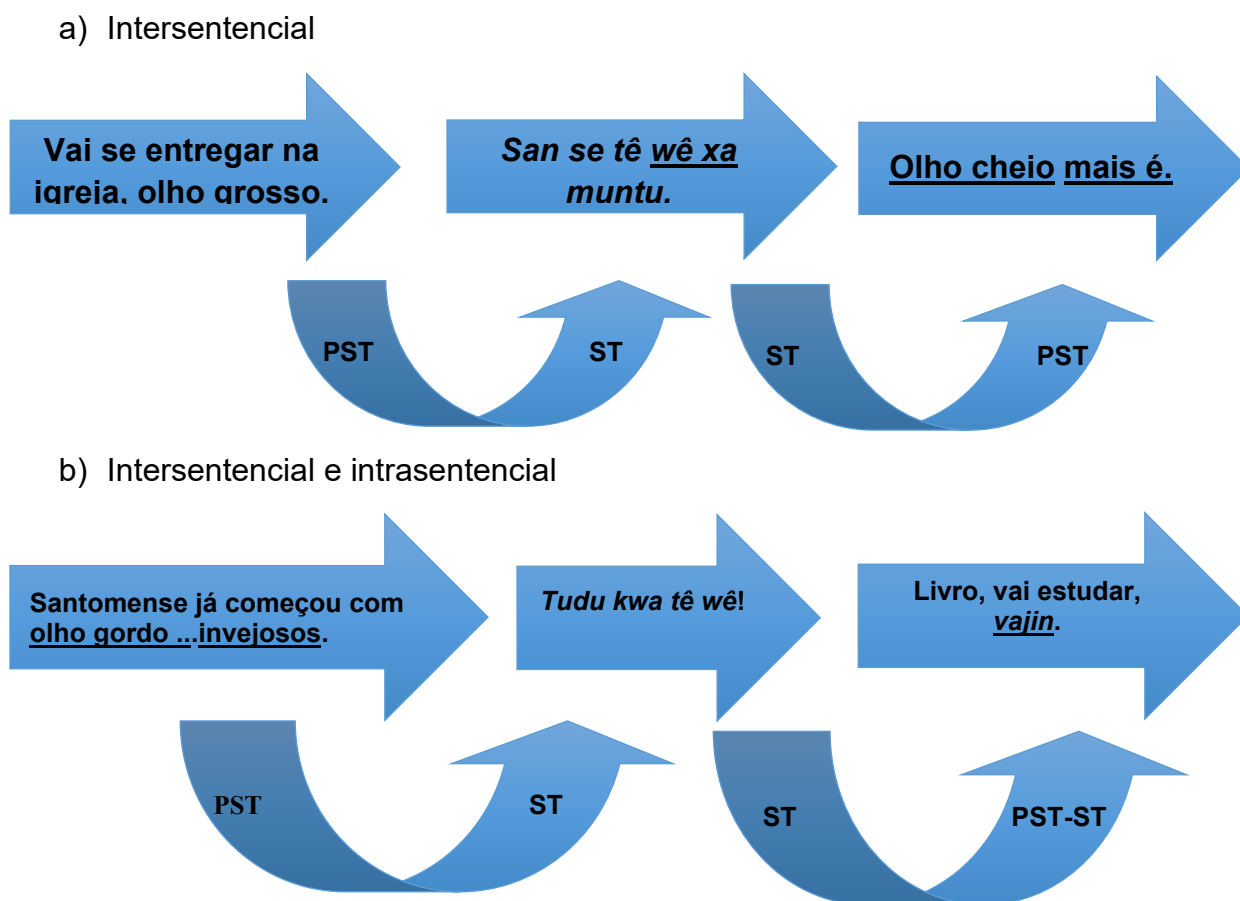
O item lexical *Pambi* não faz parte da fauna marinha nem é uma espécie de peixe; deriva do acrônimo do Programa Alimentar Mundial (PAM), conhecido no Brasil pela sigla PMA – Programa Mundial de Alimentos. Portanto, *Pambi* deriva de um processo de derivação sufixal por meio do acréscimo do sufixo “bi” ao acrônimo PAM, resultando assim em *Pambi*. Em São Tomé e Príncipe, o PAM esteve muito presente na vida dos são-tomenses, em geral, e por meio do Programa de Alimentação Escolar, que oferecia, entre os vários gêneros alimentícios, um peixe seco em postas, conforme matéria de jornal local *Tela Non* de 2011:

Realmente o nosso povo anda a dormir[.] [O] [P]into como se chama, foi o presidente do partido único e dei[x]ou o povo do jeito que dei[x]ou a comer fuba com bicho, **peixe pambi** torrrrrrrrrrrrrrado que nem a lenha, mesmo assim o povo não esqueceu das bichas [filas] e da[s] manivelas que tocamos para tirar o p[e]tróleo nas bombas. Ca[ce]tete dos policiais são-tomenses. Que raio do povo nós somos! (Tela Non, 2011) (grifo nosso)

Além dos processos de lexicalização e de gramaticalização da Figura 19 (p. 108) já descritos e analisados, apresentamos também ocorrências de *code-switching* PST-ST-PST entre as intervenções dos usuários nos comentários 5, 6 e 7 (ver Figura 19, p.108 / Figura 20a, p. 111) assim como nos comentários 10, 11 e 12 (ver Figura 19 /

Figura 20b, p. 111). Em ambos os casos (a e b), verificamos que a alternância de código ocorre com usuários diferentes, ou seja, quando alterna o usuário, também alterna o código. Na alínea b, no entanto, temos uma dupla alternância – intersentencial e intrasentencial. Essa mudança de código intrasentencial no comentário 12 – *vajin*ST [desocupado(a)^{PT}] (Figura 19, p.108 / Figura 20, p. 111) – pode ser considerada, na verdade, um empréstimo feito do ST pelo PST, pois é usado naturalmente pelos são-tomenses quando falam português.

Figura 20: Code-switching intersentencial entre falantes/usuários diferentes



Fonte: Elaborado pelo autor

Diferentemente dos itens lexicais “fulumento”, “mafeismo” e “wexaismo” (p. 104), *banho* pode ser considerado um exemplo característico de gramaticalização que, segundo Neves (2004, p. 18), “é quando uma palavra autônoma, isto é, uma forma livre, assume atribuições gramaticais” ou ainda segundo Martelotta et al. (1996, p. 12) “quando um item lexical ou construção sintática assume funções referentes à organização interna do discurso”. *Banho* é um substantivo comum que ganhou conotação específica no cenário sociopolítico são-tomense. *Banho*, que deriva do verbo “banhar”, sofreu uma discursivização e passou a designar “fenômeno de corrupção, popularmente aceito, que ocorre durante o processo eleitoral, através da compra de votos, ou seja, “doação” de dinheiro vivo, bens móveis ou imóveis em troca de votos”. Importa frisar, porém, que *banho*, no contexto político, é tipicamente utilizado na voz passiva e, geralmente, regido apenas pelo verbo “dar” – *dar banho* (ver exemplos 34 e 35) e o verbo “banhar” não é usado com a mesma conotação:

(34) Vai me dar (um) banho?

(35) Me deram um banho.

A discursivização é aqui empregada como um dos quatro processos de mudança linguística trazidos por Castilho – lexicalização, discursivização, semanticização e gramaticalização. Segundo a Proposta Funcionalista de Mudança Linguística de Castilho (2003):

A discursivização ultrapassaria a sentença como limite máximo da análise linguística. Discurso é a interação linguística em presença, discurso é conversação, organizada por um "aparato enunciativo", que inclui o locutor, o interlocutor, o assunto, e a rede de imagens que os falantes constituem a respeito deles mesmos e de suas pressupostas

posições no que se refere ao assunto. (Castilho, 2003 *apud* Fortunato, 2008, p. 1396)

(36) Onta ocê bom?

Figura 21 – Ontá ocê bom (Vava Neviense, cantor são-tomense)



Fonte: Lourenço da Silva (2021)⁴⁰

O item lexical *Onta ocê bom?* (ver Figura 21 / Quadro 10, p. 115), tem todas as palavras da língua portuguesa, duas passaram por mudança: *onta*, na última sílaba, ensurdeceu o fonema /d/ e rebaixou o fonema /a/ do pronome interrogativo “onde”, e reduziu para “on” que é a primeira sílaba. Na sequência, a primeira sílaba do verbo estar, que aparece na terceira pessoa do singular “está”, reduz-se para “tá” e, em seguida, ocorre a junção por aglutinação das formas “on” + “tá” que correspondem ao formativo

⁴⁰ Disponível em: <https://www.stp-press.st/2021/01/02/vava-neviense-melhor-cantor-do-ano-e-radio-nacional-volta-ser-a-melhor-instituicao-do-ano/>. Acesso em: 21 ago. 2025. Vava Neviense foi eleito melhor cantor do ano de 2021 pela Rádio Nacional de STP com a música “Boquitó” (Onta Ocê Bom?).

“onta”. Para Bybee (2016, p. 171), “tal redução se deve à automatização de movimentos articulatórios nessas sequências; conforme essas cadeias são repetidas, elas se tornam mais fluentes com mais sobreposição e redução de movimentos”. E “ocê” parece seguir, também em São Tomé, o processo evolutivo “vossa mercê > vosmecê > você > ocê > cê”.

Apesar de ser um item lexical estar do PST, argumentamos que “Onta ocê bom?” tem origem na expressão *Bô bô?* do Santome. *Bô bô?*ST tem duplo sentido, podendo significar literalmente “Cadê você?” (localização geográfica) ou “Onta ocê bom?” (qualidade, posse). “Onta ocê bom?”^{PST} é decalque da forma completa de *Bô bô?*ST – *Bô ku bwa bô?*ST [$\cong$ O que você tem de bom?] que, no processo de mudança linguística, sofreu a queda dos dois itens internos *ku bwa*ST [que é bom]. Prova disso é a existência de expressões semelhantes como *Kwa bô?*ST [Cadê a coisa? / O que essa coisa tem de bom?] e *Kwa ku bwa bô?*ST [O que essa coisa tem de bom?], em que a primeira é a forma evolutiva/reduzida da segunda. Porém, “Onta ocê bom?” é sempre usado com conotação negativa, significando “Você não tá com nada.” e resultou no aumento de ocorrências de casos de violência doméstica⁴¹ na sociedade são-tomense ((ver Quadro 10, p. 115).

Nos casos de banho e onta ocê bom, Camara Jr (1986 *apud* Fortunato, 2008, p. 1394) denomina de palavras lexicais e as define como “vocábulo providos de semantema, ou seja, têm traços que remetem ao ambiente biossocial extralinguístico”.

⁴¹ De acordo com as pesquisas, a expressão surgiu em 2020, virou título de músicas e temática de teatro e até entrou em discussão do Parlamento São-tomense – Assembleia Nacional de S. Tomé e Príncipe, visto que o seu uso vinha estimulando aumento de violência doméstica no país (ver Quadro 9).

Quadro 10 - Matéria sobre *Ontá ocê bom*

Deputados apelam às mulheres a se absterem de usar a expressão “ontá ocê bom” Novembro 10, 2020 Por: Manuel Dênde, Jornalista da Agência STP-Press São Tomé, 10 Nov. 2020 (STP-Press) – Um Vice-presidente da Assembleia Nacional de São Tomé e Príncipe apelou hoje, em São Tomé, as mulheres do seu País para se absterem de usar expressões populares de natureza pejorativa e que ofendem a honra dos homens diante da mulher. Levy Nazaré que falava na manhã de hoje na sessão da Assembleia Nacional (Parlamento) afirmou que o propósito do seu apelo visa desagravar o índice de violência doméstica em São Tomé e Príncipe causado, entre outras, no seio da Família. Levy Nazaré na circunstância referia-se a expressão “<i>ontá ocê bom</i>”, ou seja, “aonde é que está você bom”, expressão que este dirigente usou num contexto político numa Sessão Parlamentar anterior e que reforçou a vulgarização de cunho pejorativo contra os homens. “Eu usei-a num contexto político e hoje, sei que a expressão está a provocar nomeadamente a situações de violência doméstica”, reconheceu. Mais adiante apelou “as mulheres, por favor, para não usarem expressões deste género tendo em conta o machismo vigente no País e a minha ideia é evitar-se situações de violência doméstica no nosso País”.	O incidente entre um homem e a sua mulher, produziu-se há pouco mais de um mês, na zona sul da Ilha de São Tomé, concretamente no Distrito de Caué, quando este mal viu-se acusado em <u>língua fôrro</u>⁴² “<i>ontá ocê bom</i>”, partiu para agressão violenta causando lesões graves na mulher. Por sua vez, Maria das Neves, da Bancada do Movimento de Libertação de São Tomé e Príncipe / Partido Social Democrata (MLSTP/PSD) a par de Levy Nazaré reforçou o apelo as mulheres para situações de abstenção, denunciando que a violência doméstica “está a tomar proporções graves no País”, ao qual acresce-se a violência doméstica a situação de abuso sexual. Por seu turno, o Deputado José Miguel da Bancada do partido Acção Democrática Independente (ADI), denunciou a recente violência entre alunos numa escola básica na vila de Bombom. Defendeu na circunstância que além de medidas administrativas, nomeadamente, aplicação de Leis, é preciso envolver estruturas da juventude e das mulheres no combate aquilo que considerou de “males que afectam seriamente o futuro de gerações em São Tomé e Príncipe”. Fim/MD
---	--

Fonte: STP Press (2020)⁴³

⁴² O fato de a expressão “ontá ocê bom” estar em português e matéria afirmar que está na língua fôrro (Santome) reforça o nível de impacto do ST no PST e a dificuldade de muitos são-tomenses de delimitar a fronteira entre o PST e ST.

⁴³ DÊNDE, Manuel. Deputados apelam às mulheres a se absterem de usar a expressão “ontá ocê bom”. Disponível em: <https://www.stp-press.st/2020/11/10/deputados-apelam-as-mulheres-a-se-absterem-de-usar-a-expressao-onta-occe-bom/>. Acesso em: 10 Nov. 2020.

Além desses, há uma mudança lexical em andamento e um neologismo em ascensão no PST. Trata-se do item lexical “Stpense”, que tem sido usado em algumas comunidades virtuais de são-tomenses. Stpense, como se pode observar, é uma derivação da sigla do país – STP (São Tomé e Príncipe), que se torna um termo concorrente do são-tomense e santola, sendo este último uma gíria. O termo surge como resultado da autossensibilização e de unidade nacional para que as pessoas nativas da Ilha do Príncipe se sintam incluídas no gentílico nacional, uma vez que muitas se sentem discriminadas pelo termo oficial “são-tomense / santomense”, argumentando que faz referência apenas à ilha de São Tomé, como ilustrado a seguir:

(37) Gentílicos de São Tomé e Príncipe

- a) São Tomé → São-tomense
- b) Príncipe → Principense
- c) São Tomé e Príncipe → São-tomense ou Santomense
- d) STP + ense = STPense**

(38) Não deixemos saber como são tratadas pessoas stpenses nas Ilhas.

(39) O stpense está perdendo dinheiro comendo ouro como livro.

(40) É de salutar a ideia, iniciativa e mais posição do governo STPense em exigir esse tipo de parceria com início imediato, não quando Cuba e China quiserem daqui 10...15 anos.

(41) Excelente tarde! Conheço um Stpense que está iniciando um projeto de comércio exterior.

(42) Boa tarde, stpenses! Alguém sabe o significado de Caué, Pagué, Lembá, Lobata e Mé-Zochi?

4.2.5. Itens lexicais emprestados do Santome: fauna, flora, topônimos e projetos

Os itens lexicais apresentados nos quadros a seguir são empréstimos provenientes do Santome e não se enquadram necessariamente no processo de mudança linguística. Estão incluídos no item “Lexicalização” e não no “*Code-switching*” por defendermos que não possam ser classificados como tal. *Code-switching*, ou alternância de código, ocorre tipicamente quando o falante tem à disposição os recursos linguísticos de mais de uma língua e faz a alternância entre essas línguas por conveniência ou necessidade, de forma intencional ou espontânea. Nos exemplos listados abaixo (Quadros 11 e 12, pp. 117-118), no entanto, os falantes não dispõem desse vocabulário nos dois códigos (exceto “voador”). Esses itens lexicais só existiam no ST e foram incorporados ao português. Portanto, assim é o Português de São Tomé.

Quadro 11 - Fauna

PST/ST	Fonética	Português	Nome científico
Padê-kampu ^{PST/ST} (n.)	[pa'dɛ 'kɛpu]	Bispo-de-coroa vermelha	<i>Euplectus hordeaceus</i>
<u>Muncanha</u> ^{PST} / Munken ST (n.)	[mũ'kɛ]	Pomba-preta	<i>Aplopelia larvata simplex</i>
Vadô-panhá ^{PST/ST} (n.) ou Vadô ST e Vador ^{PST}	[va'do pã'ja]	Peixe-voador	<i>Cheilopogon melanurus</i>
Ôsôbô ^{PST/ST} (n.)	[oso'bo]	Cuco-esmeraldino	<i>Chrysococcyx cupreus insularum</i>
Fulu-fulu ^{PST/ST} (n.)	['fulu'fulu]	Atum-bonito	<i>Auxis thazard</i>
Tluki-sun-dêsu ^{44 ST} (n.)	['tluki 'sũ 'desu]	Truqui(-sun-dêsu) ^{PST}	<i>Prinia malleri</i>

Fonte: Elaborado pelo autor

⁴⁴ *Sun Dêsu*ST = Senhor Deus (≅ Truqui de Deus) – um pássaro de pequeno porte, considerado sagrado.

Quadro 12 - Flora

PST/ST	Fonética	Português	Nome científico
kundu-mwala-ve (n.)	[kũ'du 'mwala 'vɛ]	Kundudi-mwala-ve	<i>Acanthus montanus</i>
Mukumbli (n.)	[mukũ'bli]	Mukumbli.	<i>Lannea welwitschii.</i>
Fya-glêza (n.)	['fja 'gleza]	Feto / Samambaia ^{PB}	<i>Dryopteris filixmas</i>
(Folha) mikoko (n.)	[miko'ko]	Alfavaca-cravo ^{PB}	<i>Ocimum gratissimum</i>
(Folha) Salakonta (n.)	[sala'kõta]	Cana-de-jardim ^{PB}	<i>Canna bidentata</i>
Dawa (n.) ⁴⁵	['dawa]	Coco verde ^{PB}	---
Valudu (n.)	[va'ludu]	Coco seco ^{PB}	---
Xtlêlê	[tʃle'le]	Beija-flor-pequeno de S.Tomé	<i>Anabathmis newtonii</i>

Fonte: Elaborado pelo autor

Além desses itens lexicais da fauna e da flora são-tomense, a língua Santome se faz presente em todos os outros ambientes da sociedade. Apresentamos, a seguir, alguns topônimos (Quadro 13, p. 118) e projetos sociais e governamentais (Quadro 14, p. 120) de São Tomé e Príncipe cujos nomes estão em PST, ST ou em ambos. Portanto, eles também se enquadram tanto nas categorias de lexicalização e *code-switching* e foram listados separadamente para demonstrar a complexidade linguística de STP e de seus processos de mudança linguística.

Quadro 13 - Topônimos

PST/ST	ALUSTP	Português	Categoria
Bubu-Budu	=	Pedra-Pedra	Bairro
<u>N'Guembú</u> Nature Resort	<i>Ngêmbu</i>	Morcego	Resort

⁴⁵ No Português de São Tomé (PST), o “coco” recebe quatro designações de acordo com a fase em que está: *dawa*^{ST/PST} ['dawa] = coco verde; *bode*^{PST} ['bɔdʒi] ≅ quando a polpa do coco verde começa a endurecer; coco = coco seco, que ainda contém água; *valudu* [va'ludu] ≅ coco seco que já não contém água e que a polpa está se desprendendo ou já se desprendeu da casca. No Santome (ST), porém, o item lexical *kôkôndja* [ko'kõdʒa] designa tanto “coco” como o “coqueiro”. Portanto, três das quatro designações são empréstimos do ST – *dawa*, *bodje* e *valudu*.

<u>Omali</u> São Tomé (Figura 22)	=	Mar	Hotel
Vêndê Pleçu Doxi ST	<i>Vêndê Plesu Doxi</i>	Loja Preço Baixo	Loja/mercado
Pléçu Ceto	<i>Plesu Seto</i>	Preço certo	Loja/mercado
Pousada Gente d'aqui ⁴⁶	=	Ô de Casa ^{PB}	Pousada
Hotel Residencial Luxan	=	Bairro	Hotel/Bairro
Ban Bé Non Tours	<i>Ban Be Non Tours</i>	Vamos embora ^{PT} / Bora ^{PB}	Agência de Turismo
Liceu Sum Mé Xinhô	<i>Sun Me Xinhô</i>		Liceu / Colégio
Ilhéu Quitxibá	<i>Ilhéu Kitxiba</i>	Banana-prata ⁴⁷	Ilhéu
Mato Quitxibá	<i>Matu Kitxiba</i>	Banana-prata	Bairro

Fonte: Elaborado pelo autor

Figura 22 - Hotel Omali São Tomé



Fonte: Hotel Hunter (2023)⁴⁸

⁴⁶ Decalque do item lexical do Santome *Ngé d'ai éê*, usado para anunciar chegada a uma residência.

⁴⁷ O item lexical *quixibá*^{PST} admite duas grafias no Santome – *kitxiba*ST e *kixiba*ST. E a “banana-da-terra” em PST é chamada de “banana-pão”^{PST} – *bana-mpon*ST [’bana ’mpõ]. Porém, dependendo do contexto, *bana* pode se referir a qualquer espécie de banana ou bananeira.

⁴⁸ Disponível em: <https://hotelhunter.com/en/sao-tome-and-principe/sao-tome/omali-sao-tome>. Acesso em 21 mai. 2024.

Quadro 14 - Projetos sociais e ambientais

PST/ST	ALUSTP	Português	Categoria
<i>Mina Muala Non</i>	<i>Mina Mwala Non</i>	Nossas Meninas	Projeto social
<i>Zuntamon</i>	<i>Zuntamon</i>	Juntar as mãos / União	Projeto social
<i>Liqueza Téla Nón</i>	<i>Likêza Tela Non</i>	Riqueza da Nossa Terra	Projeto ambiental
<i>Omali Vida Nón</i>	<i>Omali Vida Non</i>	Mar Nossa Vida	Projeto ambiental
<i>Luxan Nón Limpo</i>	<i>Luxan Non Limpu</i>	Nosso Bairro Limpo	Projeto ambiental

Fonte: Elaborado pelo autor

Quadro 15 – Jornais digitais e/ou páginas nas mídias digitais

PST/ST	ALUSTP	Português	Categoria
Téla Non	<i>Tela Non</i>	Nossa Terra	Jornal digital
Kuá Nón	<i>Kwa Non</i>	Nossa Coisa	Página/grupo
Kê Kuá	<i>Kê Kwa</i>	O quê?	Página/grupo
Bamu flá lungua non	<i>Bamu flá lungwa non</i>	Vamos falar a nossa língua	Página/grupo
Vozes d'ôbô	<i>Vozes d'ôbô</i>	Vozes da Roça	Página/grupo

Fonte: Elaborado pelo autor

4.3. GRAMATICALIZAÇÃO

Apresentamos a seguir exemplos de gramaticalização de itens lexicais do PST conforme explanado na introdução deste capítulo, com ênfase na estrutura (sintaxe) contrapondo-se ao ST, mas ciente de que o léxico impacta a estrutura. Em outras palavras, embora o objetivo aqui seja a estrutura de sentenças no PST, nem sempre a análise do item lexical poderá ser desconsiderada. Buscamos, aqui, demonstrar por meio dos dados coletados como a gramática do PST é amplamente impactada e de diversas

formas pelo ST. Além disso, importa ressaltar que alguns exemplos apresentados no item 4.2. Lexicalização (p. 102), especialmente, “banho”, “mazé” e “onta (v)ocê bom” se enquadram igual e proporcionalmente na categorização lexical e gramatical e já foram analisados em ambas, tornando-se, portanto, desnecessária a reanálise neste item 4.3. (Gramaticalização).

(43) Uso do possessivo e pronome oblíquo como recurso enfático no PST

- a) Eu tou com minha fome.^{PST}
 1SG ESTAR PREP POS FOME
- b) Eu tou com fome de mim.^{PST}
- c) *N sa ku fomi mu.*ST
- d) Eu tou com fome.^{PT/PST}

O exemplo da alínea “d” é comum e corrente no português padrão (PT), inclusive também em São Tomé e Príncipe (STP). Entretanto, as alíneas “a” e “b” também são correntes no PST, sendo a “b” menos frequente, mas ambas têm como base a estrutura do Santome: na “a”, o possessivo “minha” preserva a ordem do PT, mas a “b” mantém a sintaxe do Santome, com o pronome pessoal do caso oblíquo tônico “de mim” funcionando como um possessivo “minha”. Importa enfatizar que *mu* no Santome pode ser tanto um possessivo, pronome pessoal átono (me) ou tônico (mim). Neste caso, o *mu* possessivo no ST se gramaticalizou no PST como forma possessiva “a”, embora não indique relação de posse/propriedade de fato, e pronome oblíquo tônico “b”.

(44) Uso do possessivo como recurso enfático no PST

a) Coisa caiu dele.^{PST}
 COISA CAIR 3SG/POS

b) *Kwa* *kê* *dê*.ST
 COISA CAIR 3SG/POS

c) A coisa caiu. (*lit.*)

O exemplo 44, tem as mesmas características do exemplo 43, onde *dê*, um pronome oblíquo átono no ST, alínea “b”, se gramaticalizou também no PST como pronome átono e preservou a estrutura gramatical do ST, alínea “a”. O sentido aqui, no entanto, não é literal - “coisa” refere-se ao órgão genital masculino. Portanto, no contexto da frase, que é o título e refrão de uma música, “Coisa caiu dele” significa “o homem tornou-se impotente”. Essa mesma frase em outro contexto de uso, no entanto, poderia ter simplesmente o sentido de “A coisa caiu”, seu sentido literal como está no exemplo 44(c). Além disso, na mesma música, observa-se a ocorrência de *code-switching* – *Oze kwa kê dê*. = Hoje a coisa caiu dele/dela^{PST}.

(45) Uso do possessivo e pronome oblíquo como reflexivo no PST

Eu tou ir embora de mim eh.^{PST}
 1SG ESTAR IR EMBORA PREP 1SG IDEO

N ska be mu ê.ST

➤ Eu tou indo embora.^{PT} / Eu vou(-me) embora.^{PT}

No caso do exemplo 45, todas as alíneas fazem parte do PST, mas a alínea “a” é característica do PST. O *mu* nesse caso é pronome oblíquo átono no ST (b) que se gramaticalizou no PST como pronome oblíquo tônico (a), mas preservou a sintaxe do

ST. Nas alíneas “c” e “d”, no entanto, ele se mantém como pronome oblíquo átono, porém com a função de pronome reflexivo “me” – Vou-me embora.

Os exemplos 46 (Minha cabeça não dói eh), 47 (Minha cabeça não abre eh.) e 48 (Cabeça vai querer abrir gente.), (pp. 124-125), são empréstimos do ST e preservam os seus constituintes no eixo sintagmático. No exemplo 46, a expressão “cabeça doer” significa “adoecer” enquanto os casos apresentados nos exemplos 47 e 48 têm sentido de “enlouquecer”. Nas três situações, observamos uma nova carga semântica transferida do ST para o PST, uma vez que a dor de cabeça desempenha a função de uma metonímia, pois indica parte de um problema de saúde maior, e “cabeça abrir” (48) insere uma nova combinação ou fraseologia no PST. Além disso, os exemplos 46 e 47 carregam uma carga sociocultural bastante relevante porque os falantes do ST usam para alertar o interlocutor que gozam de boa saúde física e mental, respectivamente, e que se forem acometidos por qualquer enfermidade, seria consequência de um feitiço feito pelo interlocutor.

Na continuidade da discussão sobre fraseologias, recorreremos às contribuições de Bally (1951 [1909]) e Faulstich (2011). Citando Bally (1951 [1909], Barbosa (2023, p. 17) define unidades ou locuções fraseológicas como associações de palavras que podem, não só, se decompor após a criação, nas quais os constituintes são relativamente livres como podem também adquirir um significado único, que não se justifica pela observação dos constituintes em separado sobre agrupamentos de lexias. Já para Faulstich (2011), Barbosa (2023, p. 17) afirma:

Esses agrupamentos são “frases parêmias” formados por um grupo vocabular coeso com restrição a alterações dos constituintes no eixo sintagmático, que alcançam estatuto de cláusula, porque têm sentido completo e desempenham papéis de forma, de ordem e de prosódia na expressão linguística. (Faulstich, 2011 *apud* Barbosa, 2023, p. 17)

Os dois exemplos em análise se enquadram exatamente nas frases parêmtias discutidas por Faulstich, pois são formados por um grupo vocabular coeso, compartilham dos mesmos constituintes no eixo sintagmático e admitem apenas duas possibilidades no eixo paradigmático – troca da forma verbal “dói” por “abre”:

(46) Minha cabeça não dói eh!^{PST}

1SG/POS cabeça NEG abrir IDEO

*Kabesa mu na ka dwê fa ê!*ST

Cabeça 1SG/POS NEG ASP doer NEG IDEO

➤ Se eu adoecer, a culpa é sua.^{PT}

(47) Minha cabeça não abre eh!^{PST}

1SG/POS cabeça NEG abrir IDEO

*Kabesa mu na ka bili fa ê!*ST

➤ Se eu enlouquecer, a culpa é sua.^{PT}

(48) Cabeça vai querer abrir gente.^{PST}

Cabeça ir querer abrir gente

*Kabeça ka mêsê bili ngê.*ST

Cabeça ASP querer abrir gente

➤ ≅ Parece que a gente vai enlouquer.^{PT}

O exemplo 49 também compartilha das características do ST e sofreu mudança linguística após a lexicalização, alterando simultaneamente a gramática do PST. O item lexical “*lcé*” resulta da aglutinação do pronome demonstrativo “isso” e a forma verbal da

terceira pessoa do singular do verbo “ser” – é. Entretanto, ocorreu apócope, com a supressão fonética do “o” do demonstrativo “isso” – **isso é > iss’e > isse > ice**.

(49) **Icé** trabalho!^{PST}

Isso é trabalho!

*Ê sa stluvisu!*ST

➤ Que bonito, hein!^{PB} / Não tem o que fazer!^{PT}

Icé trabalho!?^{PST} < Isso é trabalho!?^{PST} < Ê sa stluvisu! / Kwa se sa stluvisu!?ST

Uso do pronome oblíquo “mu”, advérbio de negação “non” e da preposição “cu” do ST no PST.

(50) **Miux!** Não **lunia** **mu** nessas coisas não **é**, coisa bonito muito.
IDEO NEG LUMIAR PRON IDEO

*Miux! Na lumia mu ni inem kwa se fa ê. Kwa glavi muntu!*ST

➤ Não coloca meu nome nessas coisas. Muito bonito!^{PT}

(51) **Taqui** **mu** aqui.
CTR (ESTAR + AQUI) PRON AQUI

*Ya mu naí!*ST / *N sa naí.*ST

➤ Estou aqui.^{PT} / Eis-me aqui.^{PT}

(52) **Deixamu** **cu** minha vida.
DEIXAR PRON COM POS VIDA

*Desa mu ku vida mu!*ST

➤ Me deixa (em paz).^{PT}

- (53) Non fazmu fala hem.
 NEG FAZER PRON FALAR HEIN.

*Na fe mu fla fa é!*ST

➤ Não me faça falar.^{PT}

- (54) Coisa aqueceu. JBJ, você chamômu?
 COISA AQUECER JBJ, PRON CHAMAR MU.

(...) *Bô sama mu!*ST

(...) Você me chamou.^{PT}

Nos exemplos 50, 51, 52, 53 e 54, verificamos o pronome pessoal átono “mu” do ST sendo usado naturalmente no PST no lugar do “me”. De fato, é uma prática muito recorrente entre os são-tomenses quando se expressam naturalmente em contextos menos formais. Nesse sentido, argumentamos que estamos diante de uma variação linguística e que “mu” é um item lexical do PST e concorrente do “me”. Além do “mu”, “ku”, no exemplo 52, é um pronome emprestado do ST que também faz parte do PST, embora seja usado com menor frequência que “mu” e por pessoas menos letradas. Desse modo, afirmamos que “ku” constitui uma variante linguística da preposição “com” no PST. Além disso, o item lexical “ku” do Santome corresponde tanto à preposição “com” e ao pronome “que” do português. Assim, em outros contextos, “ku” é usado no PST no lugar do pronome relativo “que”, sendo, por isso, concorrente do “com” e “que” no PST.

O exemplo 50 do PST contém também os ideofones “Miux” e “é”, a forma verbal “lumia” e um uso particular do advérbio “muito” no PST. O ideofone “é” é objeto de análise nos exemplos (ver exemplo 12 (p. 83) e figuras 17 (p. 105) e 23 (p. 128)) e o uso do

intensificador “muito” aqui presente foi analisado no exemplo 11 (pp. 77-80) desta tese, mas esclarecemos que “(coisa) bonito muito” é uma estrutura/expressão carregada de ironia e tem conotação contrária – “coisa feia”. Com relação ao ideofone “miux” e o verbo “lunia[r]”, argumentamos que ambos também são empréstimos do Santome e “miux” é usado para demonstrar desprezo ou desaprovação de uma ação, comportamento ou fala de outra pessoa. O Verbo “lunia[r] do português, por sua vez, deriva do latim *lumen* “luz”, e significa reluzir, emanar luz, iluminar etc. O “lunia” usado no exemplo, no entanto, é um empréstimo do Santome e significa “mencionar”, “citar”, “falar” e é esse o sentido da frase. Trata-se mais especificamente de um decalque-empréstimo do ST – *Na lunia mu*ST > Não lunia mu”^{PST}, pois o item foi emprestado e parcialmente traduzido para o PST.

(55) Manos, fala verdade, deixa de novela, há indivíduo que quando te diz tudo acabou, comida que entrou sai, cabeça abre, casa de doido toma parte, há pessoa também que se acabou ou não é irrelevante.

a) (...) casa de doido toma parte

CASA PREP DOIDO TOMAR PARTE

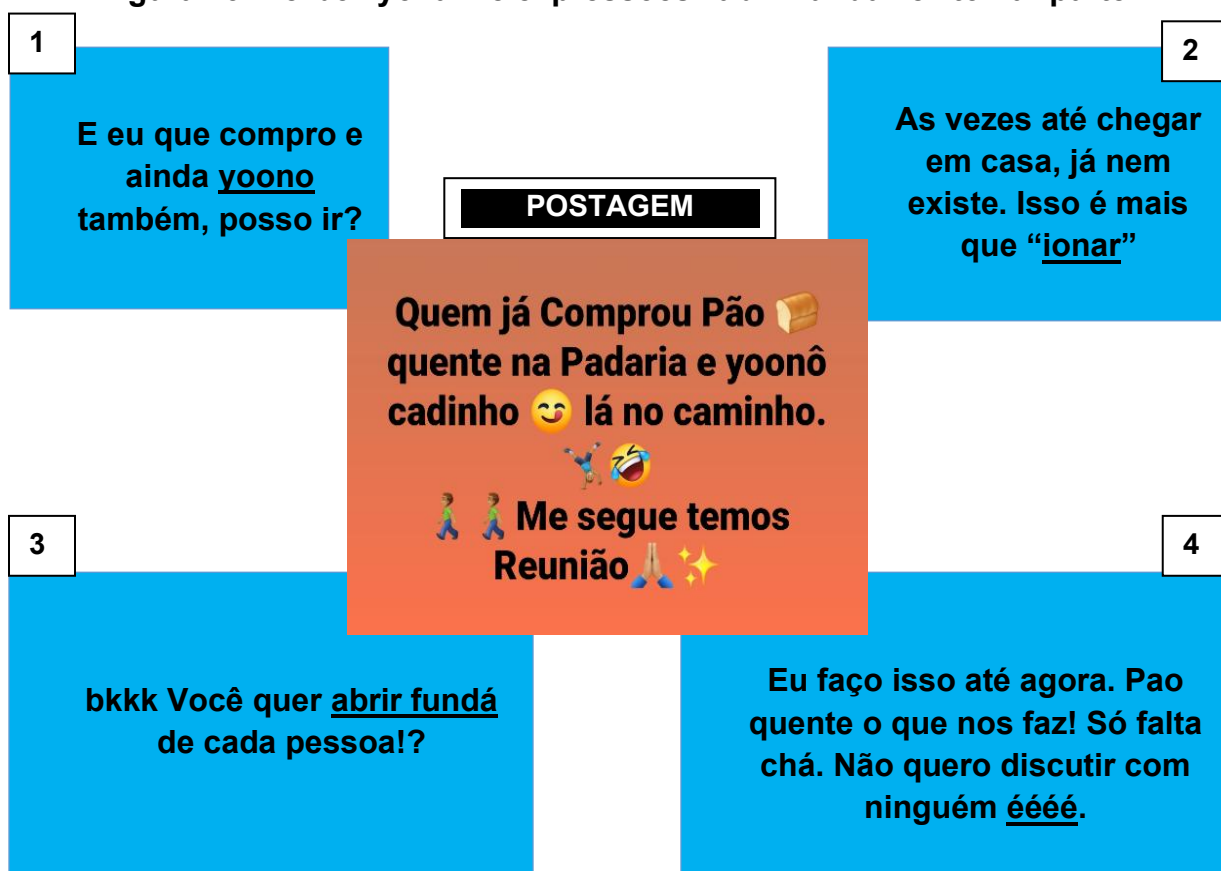
*Ke dôdô ka toma patxi!*ST

Vai parar no hospício/psiquiatria.^{PT}

O exemplo 55 contém mais uma ocorrência da expressão “cabeça abre” previamente analisada e a expressão “tomar parte”, aqui empregada, difere do uso mais comum do português – participar, intervir etc. É uma locução verbal emprestada do Santome “*toma patxi*” e significa “ir parar em” e é usado sempre com conotação metafórica. Portanto, “casa de doido toma parte”^{PST} é mais um decalque completo do

Santome *Ke dôdô ka toma patxi*ST que significa “enlouquecer”^{MET}. Salientamos, porém, que essa locução verbal é usada rotineiramente em diversos contextos, mas geralmente com conotação negativa - *ximiteli toma patxi*ST > cemitério tomou parte^{PST} = morreu^{MET}, *kadja toma patxi*ST > cadeia tomou parte^{PST} = foi preso^{MET}, *sode toma patxi*ST > polícia tomou parte^{PST} = foi para na delegacia^{MET} / a polícia foi chamada^{MET} etc.

Figura 23: Verbo “yonar” e expressões “abrir fundá” e “tomar parte”



Fonte: Elaborado pelo autor

Na Figura 23, temos uma pluralidade de processos e fenômenos linguísticos – verificamos a presença das questões foco desta pesquisa – lexicalização, gramaticalização e *code-switching*. A postagem principal introduz um novo item lexical

“yonô” (verbo), oriundo do Santome, e outro item lexical em processo de mudança linguística “cadinho” < bocadinho – aférese, supressão da primeira sílaba “bo”.

Temos ainda no comentário 3, a expressão “abrir fundá”, a criação do infinitivo do verbo “ionar” (2) e o uso do ideofone “é” no comentário 4. Todos os itens lexicais mencionados, com exceção de “cadinho”, têm origem no Santomé, mas hoje já integram o PST tanto no eixo sintagmático como no eixo paradigmático. *Yono* significa beliscar, comer pequena porção de alimento sólido, como pão, bolo etc. “Abrir fundá”, por sua vez, denota contar/revelar os segredos mais íntimos. O ideofone “é”, por seu turno, desempenha apenas a função de enfatizar a mensagem/comentário.

(56) Vêndê Pleçu Doxi ST: *vêndê* vs *bêndê*

O caso acima “*Vêndê Pleçu Doxi*” [Mercado/Loja Preço Doce] (Figura 24, p. 130) percebemos um detalhe linguístico curioso e interessante: *vêndê* e *bêndê*. Ambos os itens lexicais são do Santome e estão sendo utilizados no Santome. Porém, enquanto o substantivo é grafado com “v” (*vêndê*), o verbo grafa com “b” (*bêndê*). Além disso, o nome é uma paroxítona – *vêndê* [ˈvẽde], e o verbo admite as duas possibilidades, podendo ser tanto oxítona – *bêndê* [bẽˈde] e paroxítona – *bêndê* [ˈbẽde].

Outra questão importante a ser observada na figura é o Novo Acordo Ortográfico. Embora São Tomé e Príncipe o tenha ratificado, está evidente na fatura (Figura 24) que o Acordo não foi implementado, pois é fácil perceber os itens lexicais direção e fatura ainda sendo escritos com a consoante “c” não pronunciada, “direcção” e “factura”.

Figura 24: Nova Fatura Vêndê Pleçu Dóxi

DIRECÇÃO DOS IMPOSTOS
SÃO TOMÉ E PRÍNCIPE

NOVA FACTURA

CONHEÇA O MODELO OBRIGATÓRIO A PARTIR DE 1 DE JUNHO

VÊNDÊ PLEÇU DÓXI
NIF 109 544 000
ÁGUA GRANDE
988 88 88

CLIENTE: SAN MARIA

FACTURA
Nº FT0630122000008
01.06.2023

NIF 109 888 555

Produto	Quant.	Preço unid. (STN)	IVA%	IVA	Subtotal (STN)
SUMOL	01	12,00	15%	1,80	13,80
ÓLEO	01	55,00	15%	8,25	63,25
LEITE	01	32,00	7,5%	2,40	34,40
NATA	01	30,00	7,5%	2,25	32,25
PÃO	10	20,00	7,5%	1,50	21,50
1KG FUBÁ	01	22,00	7,5%	1,65	23,65
ESPAGUETE	01	17,00	7,5%	1,27	18,27
SABÃO	01	60,00	15%	9,00	69,00
SUBTOTAL					276,12
IVA					28,12
TOTAL					304,25

Fonte: Dire(c)ção dos Impostos de São Tomé e Príncipe⁴⁹

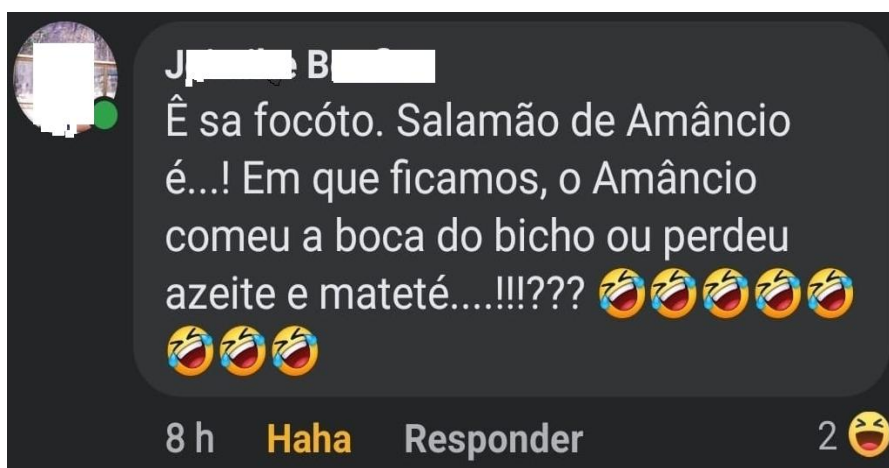
⁴⁹ Disponível em: <https://impostos.financas.gov.st/>. Acesso em 06 jun. 2023

Em suma, observamos que os exemplos analisados demonstram que o Português de São Tomé (PST), além do léxico, possui inúmeras particularidades na sua estrutura sintática como resultado do impacto do Santome. Portanto, os dados coletados mostram evidências e características que fazem do PST uma língua singular, fortemente influenciada pelo contato com o Santome.

4.4. CODE-SWITCHING

Conforme adiantado na introdução do presente capítulo, a delimitação e categorização dos dados em léxico, gramática e *code-switching* nem sempre seriam tão simples, já que muitos itens lexicais se enquadram em mais de uma categoria. Portanto, muitos itens lexicais listados para análise na lexicalização e gramaticalização também se enquadram no *code-switching*. Assim sendo, nesta seção, a ênfase é dada à alternância de código. Portanto, buscaremos agrupar e analisar ocorrências de mudança de código que se inserem mais no uso cotidiano da língua, ou seja, diálogos e postagens nas redes.

Figura 25: Postagem de JB



Fonte: Facebook (2023)

Na Figura 25 (p. 131), podemos identificar os três processos linguísticos objetos de estudo desta tese – lexicalização, gramaticalização e *code-switching*, como exposto a seguir:

(57) Ê sa fokotoST

3SG ser confusão

≅ É muita confusão.^{PB}

(58) Salamão de Amâncio é

Palma da mão PREP Amâncio IDEO

Destino do Amâncio^{PT} / Sorte do Amâncio^{PT}

*Salamon di Amancio ê...*ST > Salamão de Amâncio é...!^{PST}

(59) Em que ficamos, o Amâncio comeu a boca do bicho ou perdeu azeite e

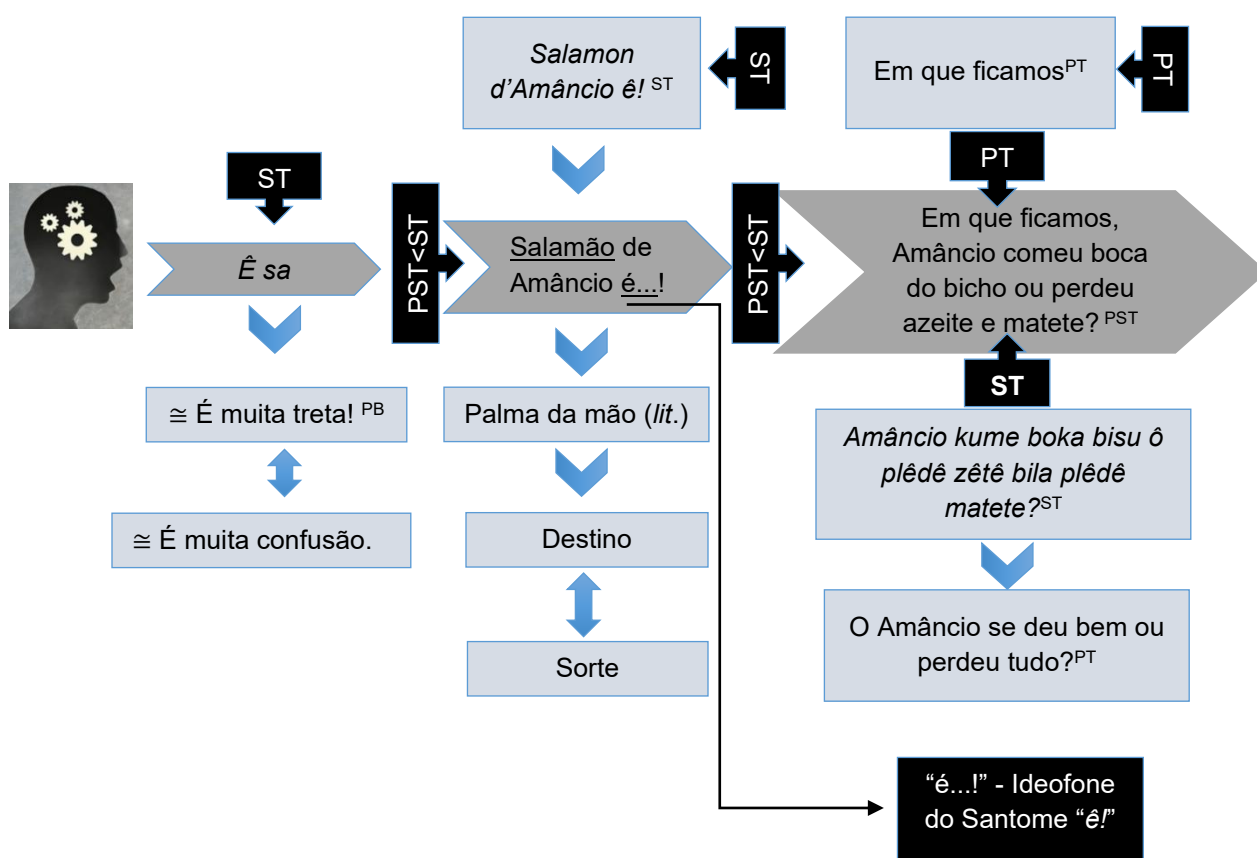
mateté!!!^{PST} [*plêdê zêdê plêdê matete*ST] = perdeu tudo^{PT}

Portanto, o falante inicia a postagem em Santome com Ê sa fokotoST, depois alterna para o PST com “Salamão de Amâncio é...!”. Essa expressão no PST é um empréstimo do ST via decalque do *Salamon di Amancio ê...*ST, incluindo o ideofone “é”^{PST}, do “ê”ST, cuja função é, nesse caso, enfatizar a situação do Amâncio. Em seguida, JB volta novamente ao PST – “Em que ficamos, o Amâncio comeu a boca do bicho ou perdeu azeite e mateté”^{PST}. Da mesma forma, “Amâncio comeu a boca do bicho ou perdeu azeite e mateté”⁵⁰ também é um decalque do Santome tanto do léxico como da sintaxe – *Amâncio kume boka bisu ô plêdê zêdê bila plêdê matete?* (ver Figura 25, p. 131).

⁵⁰ Nata residual que, na produção do azeite de palma (dêndê), se separa do produto final.

Podemos observar, nesta postagem (esquematizada na Figura 26, p. 133), o PST no seu uso cotidiano. Além da alternância de código, verificamos a ampliação do léxico do PST por meio de empréstimos de itens lexicais e da gramática do ST por meio do decalque. O nome “boca” não é tipicamente regido pelo verbo “comer” na língua portuguesa – “comer + boca”. O que temos geralmente é “comer de boca [fechada]”, por exemplo. A outra estrutura gramatical é “perder azeite e (perder) matete”, que funciona como uma unidade lexical, pois são combinações de palavras que, juntas, tem um sentido específico.

Figura 26: Processos linguísticos na Postagem de JB



Fonte: Elaborado pelo autor

Figura 27 – Postagem de PQ: *Code-switching*

Jornalista: “O Sr. foi ao quartel”?
PM: “Ao quartel?[hesitação e pausa longa] Não, não sei.Eu acho que não”.
 🙄🙄🙄🙄

and 16 others 7 comments • 1 share

N. N.
 Vc provoca... é porque a jornalista é muito persistente... chata 🙄🙄
 1d Like Reply 2

S. M.
 Mau dele...
 1d Like Reply 1

K. N.
 Olha q tu vás receber ameaças.....
 1d Like Reply 1

M. L.
 Kkkkkkkkk Kkkkkkkkkk bô minaaaaa!!!!
 1d Like Reply 1

H. T. M.
 Enfim
 23h Like Reply 1

E. B.
 Eu gostaria de saber qual substância que ele tomou esse dia para fazer ele esquecer se estava lá ou não?
 Porque eu também preciso esquecer muitos problemas desse país 🙄🙄
 Não esta fácil

and 7 others

All comments ▾

S. D.
 Esse sr que está atrás é Carlos Stoke???"
 Bo bila bi
 3h Like Reply

S. D.
 Stok conseguiu tacho???
 3h Like Reply

S. D.
 É preciso ter coragem tudo que Patrice fez
 Carlos Stok está outra vez como cão com fome a procura do pouco
 Seja homem cresce tenha dignidade com você mesmo
 3h Like Reply

Z. P.
S. D. não diga barbaridades, esta foto não é atual.
 2h Like Reply

A. Q.
S. D. esse jovem q est no fundo acho q ja na eat entre nós. Credo adis cresce na mentalidade porquê por foto antiga ainda mais qm ja se foi. N sei mas axo q nao e atual.

Fonte: Facebook, 2023

Na Figura 27 (p. 134), verifica-se que a postagem é escrita totalmente em português. Na interação nos comentários, porém, há ocorrências de *code-switching* tanto com uma manifestação isolada em frase em ST, feita por ML – *kkk Bô minaaaaa!!!*, como um comentário de SD, fazendo ele mesmo uma espécie de dupla alternância de código. Considera-se dupla porque o falante de ST alterna o código português-santome em relação à postagem original assim como uma automudança de código em relação a ele próprio – “Esse sr que está atrás é Carlos Stoke???^{PST} *Bo bila bi*.ST”

(60) *Bô mina*.ST

2SG menina!

➤ ≅ Você é terrível! / Você não vale nada!^{PT}

(61) *Bô bila bi*.ST

2SG voltar vir

➤ ≅ Você (veio) de novo! / Você é teimosa! / Você adora criar confusão.^{PT}

Figura 28 – Postagem do GS: *Code-switching*



Fonte: Facebook, 2023

A postagem do GS, Figura 28 (p. 136), é feita totalmente em português, mas temos também ocorrências de *code-switching* nos comentários, praticamente nos mesmos moldes da Figura 27 (p. 134). Observa-se um comentário do usuário DM fazendo dupla alternância de código, português-santome-português, pois ele inicia a postagem em português (mesmo código da postagem original) – “Pois é” –, muda em

seguida para ST – *Clêndêço padê cluço*ST ⁵¹– e muda novamente para PST – “Ele interrogou o Arlecio já morto [...]”. E os dois comentários seguintes alternam entre o ST – *kkkkk sa cada cua* ⁵² (de AS), e português – “Acho que é *brucho para falar com os mortos.” (feito por FL) e os demais seguem em PT.

(62) *Klêndêsu padê klusu*.ST

Crer em Deus padre cruz

➤ Credo! / Deus do Céu!^{PT}

(63) *Sa kada kwa*!ST

Ser cada coisa

➤ É cada uma!^{PB} / É cada coisa que acontece!^{PT}

No exemplo 62, verificamos também uma mudança linguística da expressão completa *Klê ni dêsu padê klusu*ST em curso no Santome. O item lexical *Klêndêsu* é resultado do processo de composição por aglutinação das palavras *Klê ni dêsu*. Nesse processo, porém, ocorreu apócope, com a supressão fonética do “i” na preposição “*ni*” – *Klê* (verbo) + *ni* (preposição) + *klusu* (nome). Além disso, verificamos também uma espécie de derivação imprópria, uma vez que embora *klêndêsu* derive da aglutinação de três categorias gramaticais – Verbo + Preposição + Nome, não resultou em uma classe gramatical específica, não tem sentido completo sozinho, precisa ser utilizado com os demais constituintes da interjeição, mas é um novo item lexical. Estamos, portanto,

⁵¹ *Klêndêsu Padê Klusu*!ST – ortografia segundo o ALUSTP

⁵² *Sa kada kwa*!ST – ortografia segundo o ALUSTP

diante tanto da lexicalização – novo item lexical – como da gramaticalização, pois houve mudança na estrutura sintática do Santome e, por conseguinte, do PST.

No exemplo 63 (p. 137), verificamos também uma mudança linguística em curso no Santome. A expressão completa é *Ê sa kada kwa!*ST. Nesse caso, os dados coletados mostram uma tendência para a queda do “Ê” (3SG), à semelhança do que ocorreu no português do Brasil (PB), se tornando uma frase com sujeito inexistente – *Sa kada kwa!* / “É cada uma!”^{PB}. Porém, no PB, já podemos perceber também uma tendência para a queda da forma verbal “É” em eventos de fala, de “É cada uma!” para “Cada uma!” E essa talvez seja a próxima etapa também no Santome. Nesse contexto, ressaltamos a tese de Bybee (2020 [2015]):

Todas as línguas mudam de maneira igual. Uma vez que os usuários das línguas do mundo afora operam com os mesmos processos mentais e usam a comunicação para os mesmos fins, ou fins muito semelhantes, as mudanças que emergem nas línguas desde o Alasca até a Zâmbia se encaixam nas mesmas categorias de mudanças encontradas no inglês e no francês. (Bybee, 2020, p. 32)

4.4.1. Provérbios ou ditados populares no PST

O uso de provérbios ou ditados populares, no PST, ocorre com grande frequência com recurso ao *code-switching*. Esse fenômeno natural de uso linguístico na sociedade são-tomense em geral se dá de duas formas: (1) diálogos ou discursos em PST com *code-switching* interlingual com a língua Santome, exemplificados nos registros 64, 65, 66, 67, 68 e 69-C1/C2, 70-C1, 71-C2 e 72, e (2) outras vezes através da tradução literal desses ditados do ST para PST (por meio do decalque), como mostram as ocorrências 69-C3/C4, 70-C2, 71-C1/C3⁵³, a seguir:

⁵³ A letra “C” em C1, C2 e C3 refere-se aos comentários 1, 2 e 3 das respectivas postagens – exemplo 69, 70.

(64) E como a vida não só festas, eu já cá estou para cumprir com o dever e fazer com que ninguém me aponte o dedo futuramente. ***Tlabá só ka cada tê.***ST Bom dia a todos os guerreiros. (postagem)

➤ ***Tlabá só ka cada tê***ST $\cong$ O trabalho é a fonte de todas as riquezas.^{PT}

(65) Aqueceuuuu...***Pontopé da ni giba eh...***ST A verdade só é verdade até o momento que está acontecer. Quem não sabe beber não pegue no copo. Anda, Bobi. (postagem)

➤ ***Pontopé da ni giba eh***ST $\cong$ Tocaram na ferida.^{PT}

(66) Cadê o casal que dizia “nada vai fazer eu desistir de você?” Quando estamos apaixonados, falamos à toa, yhaaa! (postagem)

Comentário: ***Amôlê sa poto.***ST

➤ ***Amôlê sa poto.***ST Amor está na porta. (*lit.*) $\cong$ O amor está no ar.^{PT}

(67) Vida em Portugal: salário mínimo 820 euros; Renda: 500 euros; energia e água: 100 euros; alimentação: 180 euros; passe: 40 euros. (postagem)

Comentário: ***Bi pia kwa ku wê bô, Anita!***ST

➤ ***Bi pia kwa ku wê bô, Anita!***ST $\cong$ Vem ver com os seus próprios olhos. ^{PT}

(68) Eu não quero assustar ninguém, mas por quê que a bíblia termina na página 2024. (postagem)

Comentário: ***Kakaê!*** (Ideofone)

➤ ***Kakaê!***ST $\cong$ Cruzes! Vixe!^{PT}

(69) Eu já vi tudo nessa vida, mas padaria sem pão é exagero já. (postagem)⁵⁴

Comentário 1: Kei! (IDEO)

Comentário 2: Awó! (IDEO)

Comentário 3: Gente tá nele já! (decalque)

Comentário 4: Mas pronto, como vamos fazer? Nós tá nele já. (decalque)

Comentário 5: Ê vô morre de mim!

Comentário 6: Ê vô morre de mim ô! (IDEO)

Comentário 7: Nom morre não!

Comentário 8: É vida éeee! (IDEO)

Comentário 9: São situações!

Comentário 10: Situações é! (IDEO)

Comentário 11: Enfim é! (IDEO)

(70) Primeiro Ministro defendeu as aulas e a juventude durante a campanha para ganhar votos e agora não tem solução para os mesmos. (postagem)

Comentário 1: **Kwa kwa kwa!**⁵⁵ Peixe morre pela boca. **Kwa ska bi!**ST

Comentário 2: Peixe morre pela boca. **Aqui é mundo.** Ruuuu! (decalque)

➤ ***Kwa ska bi!***ST $\cong$ **O que dele está guardado.**^{PT} (sentido negativo)

➤ ***Aqui é mundo!***^{PST} $\leftarrow$ ***Nai sa mundu.***ST $\cong$ **O que dele está guardado.**^{PT}
(sentido negativo) / **Aqui se faz aqui se paga.**^{PT}

⁵⁴ Ver Figura 29 (p. 143)

⁵⁵ Onomatopeia de gargalhada, risos (empréstimo do Santome).

(71) Parabéns ao pai do povo santomense! (postagem)

Comentário 1: Eu conheço bem o meu pai desde que nasci, não **lumia** povo, ok? (cf. exemplo 50, p.125 e 127)

Comentário 2: Filme desse gajo: O criminoso em fuga. Matou nossos irmãos, fugiu e pretende enganar o povo de novo, **ya bô tê tema**...ST

Comentário 3: **Dente aberto**⁵⁶ nos fez vergonha; faltou-lhe humildade e amor ao povo dessa vez.

➤ **Ya bô tê tema**!ST ≅ **Você é teimoso.**^{PT}

➤ **Dente aberto**^{PST} ← ***Dentxi betu***.ST ≅ **Boca aberta (lit.) / Sorriso falso. / Pilantra.**^{PT}

(72) **Um gajo** procura viver uma vida regrada, não come isso e também não bebe aquilo, não dá peto-peto, e de repente **pôh, bliquidi piiii**.ST (postagem)

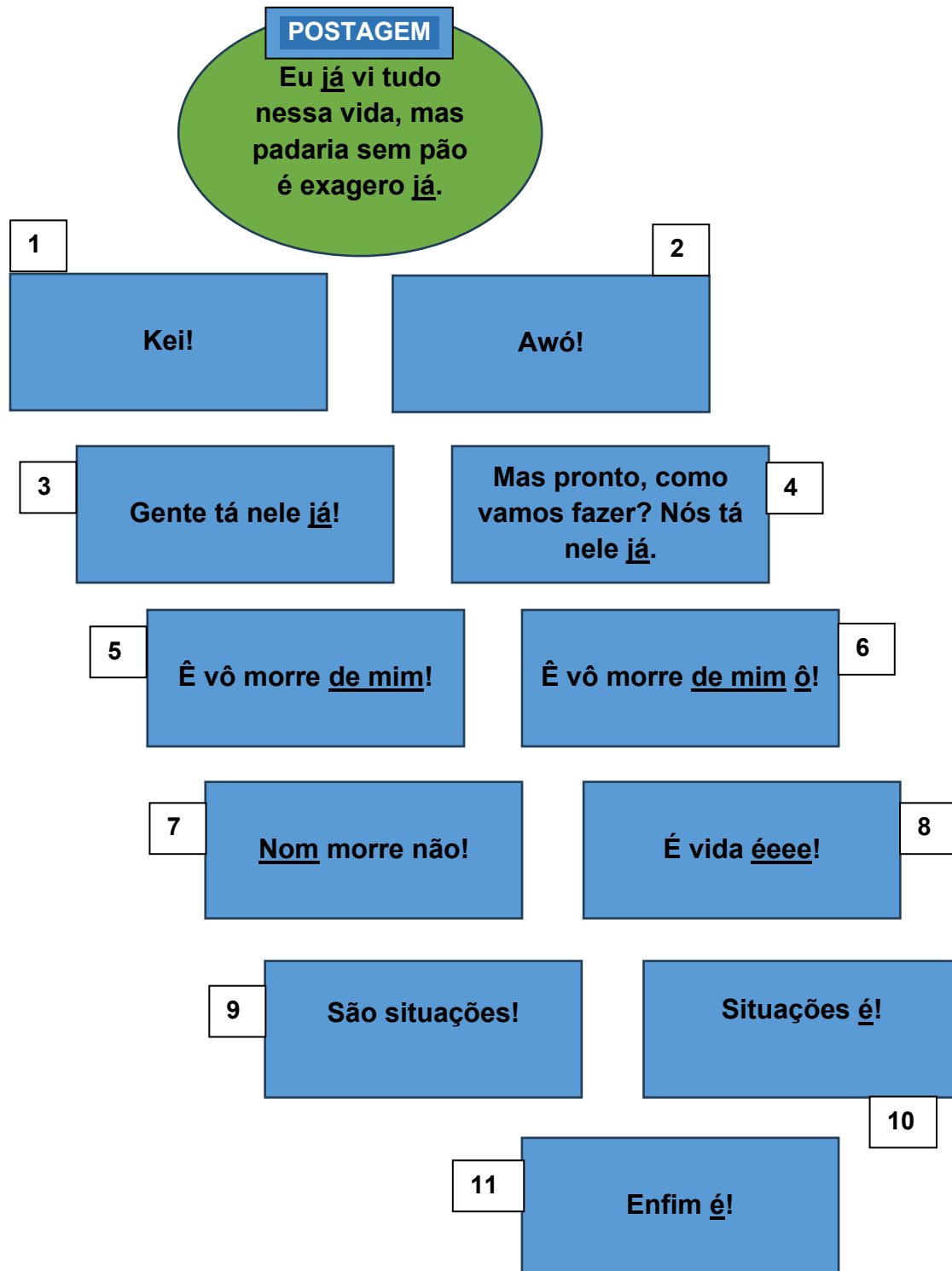
O exemplo 72 apresenta mais uma particularidade do PST – “um gajo”, pronunciado e escrito muitas vezes como “gás”. “Gajo” denota uma pessoa qualquer de sexo masculino cujo nome não se sabe ou não se quer mencionar – cara, fulano, sujeito etc. Paralelamente, usa-se “gaja”, nome comum, para alguém do sexo feminino cujo nome não se sabe ou não se quer mencionar. No contexto de São Tomé, informalmente,

⁵⁶ Apelido de um ex-primeiro-ministro de São Tomé e Príncipe, que se transformou no título de duas músicas do falecido cantor são-tomense Pêpê Lima (Pedro Lima). As músicas fazem críticas ao seu comportamento como político são-tomense bon-vivant e estão disponíveis no YouTube: LIMA, Pepê. **Sun Dentxi Bétu** disponível em: <https://www.youtube.com/watch?v=yze8NYD-qrY>. Acesso em: 18 set. 2025. E LIMA, Pepê. **Dentxi Bétu 2**. Disponível em: <https://www.youtube.com/watch?v=2el4MrA2aGw>. Acesso em: 18 set. 2025.

“um gajo” é um pronome pessoal da primeira pessoa do plural equivalente ao “a gente” e “nós”. No exemplo em análise, “um gajo” é usado como pronome pessoal do caso reto (sujeito). Contudo, “um gajo” pode ser usado também como pronome oblíquo átono (objeto direto: “Ele convidou um gajo para a festa.”) e pronome oblíquo tônico (objeto indireto: “Você ligou para um gajo?” Em contextos específicos, “um gajo” funciona também como pronome pessoal da primeira pessoa do singular – eu (“Um gajo faz tudo certo, mesmo assim dá confusão.”^{PST} - “Eu/a gente faz tudo certo, mesmo assim dá confusão.”^{PT})), me (“Você tá a faltar um gajo respeito.”^{PST} - “Você tá me desrespeitando.”^{PB}).

Ainda no âmbito do exemplo 72, temos o ideofone **pôh, bliquidi pii**ST oriundo do Santome e, hoje, também parte integrante do Português de São Tomé (e Príncipe). **Pôh, bliquidi pii**ST é um ideofone onomatopeico cujo primeiro item lexical “*pôh*” é geralmente usado como “bô” – *Bô bigidi pi!* – no qual *pôh/bô* remete a uma pancada/paulada, *bligidi*, à queda, e *pi* se refere à morte. Esse ideofone pode ser usado de forma literal ou metafórica e o exemplo, em questão, se insere num contexto mais amplo da sociedade são-tomense, e mundial por extensão, em que é cada vez mais frequente pessoas mais jovens e aparentemente saudáveis estão falecendo em decorrência de AVC, infarto etc. No caso concreto, a postagem se refere a um jovem são-tomense que sofreu um infarto, foi internado e faleceu. O mais interessante desse ideofone ainda é que os três itens lexicais que o compõem são totalmente independentes, podem ser e são usados isoladamente. O ideofone *pôh* no PST/ST é geralmente usado com sentido de tiro de arma de fogo, o “ííí” do “*pííí*” representa apenas ênfase no ideofone.

Figura 29 – Advérbio “já” no PST e Ideofones “é” e “ô”



Fonte: Elaborado pelo autor

Na Figura 29 (p. 143): verificamos ocorrência da lexicalização, gramaticalização e *code-switching* simultaneamente na postagem e nos vários comentários que se seguiram. A postagem principal apresenta o uso do advérbio “já” em ambas as orações – “Eu já (1) vi tudo nessa vida, mas padaria sem pão é exagero já (2).

Antes de concentrarmos na análise, seria importante tecer algumas palavras sobre os advérbios. Em linhas gerais, Almeida (2009, p. 191) define advérbio como “a palavra que modifica o sentido de um verbo, acrescentando-lhe uma circunstância (tempo, modo, lugar, causa, intensidade etc.). ou intensifica o significado de um adjetivo ou de outro advérbio. Pode também modificar o sentido de uma frase toda”, conforme exemplos abaixo:

- a) Aquele vereador fala bem. (modificando um verbo)
- b) Aquela garota é bastante magra. (intensificando um adjetivo)
- c) Essa professora fala muito alto. (intensificando outro advérbio)
- d) Felizmente, os bombeiros chegaram rapidamente. (modificando toda a frase)

Os advérbios podem ocupar posições sintáticas diversas dentro da frase, dependendo do próprio advérbio em questão, da intenção comunicativa e diversos outros fatores. No caso do advérbio de tempo “já”, ele geralmente precede o verbo – Eu já viajei o mundo; O Tomé já se levantou; A Marta já concluiu o seu curso de medicina; O projeto já está em andamento; Essa propriedade já tem dono(a) etc.

No exemplo em análise, no entanto, o advérbio “já” é colocado tanto na posição anterior como na posição posterior ao verbo – “Eu já (1) vi tudo nessa vida, mas padaria sem pão é exagero já”. E a colocação do advérbio em questão na posição posterior ao verbo também ocorre com os comentários. Argumentamos que essa é uma característica

Além das questões apresentadas, o PST também tem o pronome “a gente” que concorre com “nós”. Porém, está em processo de mudança constante, já perdeu o artigo “a” (aférese – supressão de fonema no início da palavra) como mostra o exemplo 75 (Comentário 3 – Figura 29 (p. 143)) e, informalmente, passou por metaplasmos por supressão de sons. “Gente” passou por apócope - supressão de fonema no fim da palavra: a gente > gente > gen. Exemplo: Gen vê só.^{PST}, que também é um empréstimo do Santome “*Non ka bê non so*”ST e significa “A gente se vê por aí”^{PT}.

No comentário 7 (p. 143), temos um advérbio de negação “nom” (= não) comumente usado no PST. Nesse caso, porém, não foi possível determinar a sua origem, pois dos advérbios de negação do Santome – *ino*, *na*, *fa*, *fô*, *nê* – nenhum é igual ao “non”, embora tenhamos o “*ino*” [’ino], que não ocorre nessa posição. Mesmo se ampliarmos o espaço e a localização geográfica de São Tomé e Príncipe e considerarmos o Golfo da Guiné como uma região onde a maior parte dos países seja francófono, o “*non*” do francês também não ocorre nessa posição. Consequentemente, acreditamos que seja resultado de mudança fonética do “não”.

Por fim, temos alguns ideofones na mesma postagem (p. 143). Há ocorrências de ideofones em diversos comentários, pois representam parte importante do PST. Os ideofones “é” e “ô” de todos os comentários são originários do ST e são usados sempre no final da frase. Além disso, não podem ser empregados isoladamente, exceto no caso de “Tchau ê! / Ê!” e “Fidel ô!” / Ô! (ver item 2.2. Gramaticalização – exemplos 12 e 13 (p. 83))” e desempenham apenas a função de ênfase nesses comentários. Já os ideofones *Kei!* (comentário 1) e *Awo!* (Comentário 2 – Figura 29), diferentes dos dois mencionados anteriormente, ocupam status diferente na gramática do PST. “Kei!” e “Awó!” são ideofones que denotam reações emotivas negativas – tristeza – e são usados sempre

sozinhos. As ocorrências do pronome oblíquo “de mim” nos comentários 5 e 6 (p. 143) já foram objetos de análise em casos semelhantes anteriormente apresentados nesta tese (ver exemplos 43, 44 e 45, pp. 121-122) o que dispensa uma reanálise aqui, mas essas ocorrências demonstram e reforçam sua frequência de uso rotineira no PST.

4.4.2. Ideofones no Português de São Tomé

Além dos itens lexicais, estruturas linguísticas, ocorrências de *code-switching* e alguns ideofones apresentados e analisados até aqui, faz-se necessário retomarmos os ideofones de forma mais ampla e particular pela enorme relevância que representam no Português de São Tomé e na realidade extralinguística de São Tomé e Príncipe de forma geral. A ocorrência de ideofones (assim como ditados populares e expressões idiomáticas) no PST é uma característica desse ambiente multilíngue, pois os nativos das ilhas de São Tomé e Príncipe recorrem natural e frequentemente ao *code-switching*, alternando do PST para Santome quando necessitam fazer uso desses recursos linguísticos. Vale ressaltar, porém, que o oposto também acontece, embora com menor frequência (recorrem ao PST quando estão se expressando no Santome para uso de alguns provérbios, expressões idiomáticas e ideofones originais do português).

Assim sendo, o que verificamos no uso prático nessa comunidade de fala frequentemente é a ocorrência, quase que simultânea, dos três processos/fenômenos linguísticos discutidos nesta tese – lexicalização, gramaticalização e *code-switching*. Embora os ideofones sejam passíveis de análises em qualquer um dos ou nos três processos/fenômenos linguísticos ao mesmo tempo, decidimos fazê-lo aqui (no item 4.4.

Code-switching) por ocorrer mais no contexto de *code-switching* ou em decorrência do uso.

É nessa perspectiva que recorreremos novamente a Dik (1989, p. 6) quando argumenta que “as línguas naturais são um instrumento utilizado para fins de comunicação e não faz muito sentido analisar as suas propriedades sem considerar os usos e funções. O sistema subjacente à construção de expressões linguísticas é um sistema funcional e deve sempre ser estudado observando as regras, princípios e estratégias que determinam seu uso comunicativo natural”.

E sobre ideofones, Araújo (2009, p. 23) afirma que “são um fenômeno linguístico complexo para o qual existem várias definições e nenhuma conseguiu abranger completamente sua heterogeneidade”. O autor ressalta que as tentativas de definição mais bem-sucedidas buscam enquadrar suas múltiplas características em grupos prototípicos, frequentemente restringindo-os a grupos de línguas ou áreas linguísticas específicas. Segundo Araujo, foi Doke (1935) quem apresentou a definição clássica para ideofone:

Uma representação vívida de uma ideia através do som. Uma palavra, comumente onomatopaica, que descreve um predicado, um qualificativo ou um advérbio em relação ao seu modo, cor, som, cheiro, ação, estado ou sua intensidade. (...) Deve-se apontar que geralmente, regras especiais de duração, tom e acento, aplicáveis a formas gramaticais ordinárias, diferem consideravelmente no caso dos ideofones’. (Doke, 1935 *apud* Araujo, 2009, p. 24)

Ainda de acordo com Araújo, Bartens (2000, pp. 19-20) criou uma tipologia interlinguística para distinguir três tipos de ideofones: ideofones intensificadores ou partículas exclusivas; Ideofones empregados em construções quotativas com ou sem um auxiliar (ideofones onomatopeicos, podendo expressar velocidade de um movimento

também); e, Ideofones com um *significado* independente (substantivos, verbos, adjetivos ou advérbios das línguas europeias). Em suma, podemos afirmar que um ideofone é uma palavra ou expressão que busca evocar uma ideia sensorial (som, visão, cheiro, intensidade, movimento, estado) de forma vívida, muitas vezes pela própria sonoridade e estrutura fonológica particular, funcionando frequentemente como um intensificador. Embora apresentemos diversos tipos de ideofones, isto é, ideofones com funções diversas, é essa característica intensificadora que destacamos nos ideofones do Santome e, conseqüentemente, do Português de São Tomé (PST), listados no quadro abaixo:

Quadro 16 – Ideofones do PST/ST

PST	SANTOME	LITERAL	USO/SIGNIFICADO
IDEOFONES			
<i>Guadá bô! / Deixa você!</i>	<i>Gwadá bô!</i> [gwa'da bo]	Espera você!	O seu tá guardado. (conotação negativa)
<i>Tongana bô!</i>	<i>Tongana bô!</i> [ton'gana bo]	-	Foi por pouco! (conotação negativa)
Punda ⁵⁷ Klixtu!	<i>Punda Klixtu!</i> ['pūda 'klixtu]	Pelo amor de Cristo!	≡ Pelo amor de Deus!
Adêwakongo!	<i>Adêwa-kongo!</i> [a'dewa 'kōgo]	Adeus!	≅ Já era!
Você fala vonvon demais.	<i>Fla von-von.</i> ['fla 'võ'võ]	À toa	Você fala à toa demais. / Fofoca demais!
Cuá beba!	<i>Kwa beba!</i> ['kwa 'bɛba]	Coisa de barba	≅ Uau! (conotação positiva) / Vixe! (conotação negativa) Depende do contexto e da entonação

⁵⁷ *Punda* ST pode significar pelo(a), porque, porquê, por que, por quê, por causa de.

Capotó!	<i>Kapotó!</i> [ka po'to]	-	≅ aprovação / surpresa
Zôplô d'ubuê	<i>Zôplô d'ubwê</i> ['zoplo d'ubwe]	Moviment o do corpo	≅ Simulação ^{LIT.} / Blefe ^{MET}
Ca piá!	<i>Ka pya!</i> [ka'pja]	Vai vendo!	≅ Que porra é essa! (ironia, deboche)
Caca!	<i>Kaka!</i> ['kaka]	-	≅ Caramba! Conotação positiva ou negativa (contexto de uso) – surpresa
Cacaê!	<i>Kakaê!</i> ['kaka e]	-	Conotação negativa
Cacaô!	<i>Kakaô!</i> ['kaka o]	Eita! Taporra!	Conotação positiva ou negativa (contexto de uso) – surpresa
Oh caca!	<i>Oh kaka!</i> [o 'kaka]	Eita! Taporra!	Conotação positiva ou negativa (contexto de uso) – surpresa
Oh cacaô!	<i>Oh kakaô!</i> [o 'kaka o]	Eita! Taporra!	Conotação positiva ou negativa (contexto de uso) – surpresa enfática
Caqui!	<i>kaki</i> [ka'ki]	Caramba!	Conotação negativa
Quei!	<i>Kêy!</i> ['kei]	-	Conotação negativa
Quê ô!	<i>Kê ô!</i> ['ke o]	-	Conotação negativa com ênfase
Quêêê!	<i>Kêêê!</i> ['keee]	-	Conotação quase sempre negativa com dó (ênfase)
Auó!	<i>Awo!</i> [a'wo]	-	Conotação quase sempre negativa
Auó ê!	<i>Awo ê!</i> [a'wo e]	-	Conotação quase sempre negativa com dó (ênfase)

Auo ô!	<i>Awo ô!</i> [a'wɔ o]	-	Conotação quase sempre negativa com choro (ênfase)
Chê!	<i>Xê!</i> ['ʃe]	-	Conotação positiva ou negativa (contexto de uso) – surpresa/espanto
Kidalê!	<i>Kidalê!</i> [kida'le]	-	Socorro! / Ai meu deus! /Boa! / Aêêê! (contexto de uso)
Kidalê ê!	<i>Kidalê ê!</i> [kida'le e]	-	Socorro! / Ai meu deus! (conotação geralmente negativa)
Kidalê ô!	<i>Kidalê ô!</i> [kida'le o]	-	Socorro! / Ai meu deus! (conotação geralmente negativa enfática)
Só de graça! / Duji-duji	<i>Duji-duji</i> ['duʒi' duʒi]	-	Sem necessidade / sem motivo
Verde ⁵⁸ tatali / Fili tatali	<i>Fili tatali</i> ['fili 'tatali]	-	Muito verde / Muito criança ^{MET} / Muito imaturo(a) ^{MET}
Fili petepete	<i>Fili petepete</i> ['fili pɛ'tɛpɛ'tɛ]	-	Muito verde (frutas(os))
Simples tatali	<i>Ximpli tatali</i> ['ʃipli 'tatali]	-	Completamente sem sal (comida)
Cala boca pipipi	<i>Kaboka pipipi</i> [ka'boka pipi'pi]	-	Calar a boca completamente (não dizer nada)
Rico sonosono	<i>Liku sonosono</i> ['liku sɔ'nosɔ'no]	-	Podre de rico/ Riquíssimo

⁵⁸ No Português de São Tomé (PST), existem cinco designações para as frutas de forma geral de acordo com a fase em que estão: *verde*^{PST}; criado(a), *pontudo*^{PST} [põ'tadu] ≅ quase maduro(a), de vez; maduro = maduro; melado [mɛ'ladu] ≅ extremamente maduro(a) e/ou começando a apodrecer. Geralmente não é mais consumido e/ou é deixado para os morcegos, quando ainda estão na árvore. Vale frisar que “verde” não se refere apenas a cor da fruta (que é a cor predominante das frutas no estágio inicial), mas a um estado em que a fruta ainda não pode ser consumida. “Verde” faz contraponto com o item lexical “criado(a)”, que indica que a fruta já está no ponto para consumo, mas ainda não está “pontada” (“de vez”). Além disso, no PST, “de vez” não é usado como locução adjetiva para qualificar o estado ou estágio das frutas, mas apenas como locução adverbial com sentido de “definitivamente” e “para sempre”.(Ex: O meu pai foi embora para Portugal de vez. = O meu pai viajou para Portugal e nunca mais retornou a São Tomé).

Vungunu-vangana ^{NOM} “Você está no vungunu- vangana.”	<i>Vungunu-vangana</i> [vũgu'nu vëga'na] <i>Bô sa ni vungunu- vangana.</i>	-	≅ Cambaleiar / Altos e baixos / vida instável ≅ Você está cambaleando na vida.
Preto lululu	<i>Pletu lululu</i> ['pletu lulu'lu]	-	Muito preto(a) Muito escuro / Completamente escuro
Vlêmê bababa	<i>Vlêmê bababa</i> [vlë'me baba'ba]	-	Muito vermelho
Frio kôkôkô	<i>Fyô kôkôkô</i> ['fjo koko'ko]		Muito frio
Fina lekeleke	<i>Fina lekeleke</i> ['fina lë'këlë'kë]		Muito bom / Muito bem / Excelente
Fina lelele	<i>Fina lelele</i> ['fina lëlë'lë]		Muito bom / Muito bem / Excelente
Seco klakata	<i>Seku klakata</i> ['sëku klaka'ta]		Completamente seco
Branco fenené	<i>Blanku fenené</i> ['blëku fënë'në]		Muito branco(a) / Branquíssimo(a)
Branco fenefené	<i>Blanku fenefené</i> ['blëku fënë'fënë]		Muito branco(a)/ Branquíssimo(a)
Escuro dim	<i>Kulu dīīī</i> ['kulu 'dīīī]		Muito escuro / Completamente escuro
Velho ketekete / Ve ketekete	<i>Ve ketekete</i> ['vë këtë'tëkëtë]		Muito velho / Velhinho
Molhado potopoto	<i>Monhadu potopoto</i> [mõ'jadu pö'töpö'tö]		Encharcado / Completamente molhado
Quente zuzuzu	<i>Këntxi zuzuzu</i> ['këtʃi zuzu'zu]		Muito quente
Baklaga!	<i>Baklaga!</i> [ba'kla'ga]		(Acertar) em cheio (conotação positiva)
Quebrar winiwini	<i>Kebla winiwini</i> [kë'bla wi'niwi'ni]		Fragmentar
Cortar winiwini	<i>Kota winiwini</i>		Fragmentar

	[kɔ'ta wi'niwi'ni]		
Nu tatali	<i>U'nu tata'li</i> [u'nu 'tatali]		Completamente nu
Suzu kotokoto	<i>Suzu kotokoto</i> [Suzu kɔ'tɔkɔ'tɔ]		Muito sujo/ Imundo
Danado kotokoto	<i>Danadu kotokoto</i> [Dana'du kɔ'tɔkɔ'tɔ]		Muito traquino / Muito assanhado ^{MET}

Fonte: Elaborado pelo autor

4.5. SÍNTESE DO CAPÍTULO

A maioria dos exemplos extraídos dos dados coletados apresentam ocorrências de *code-switching* português-santome. Verificamos que são mais comuns dois tipos de ocorrências – *code-switching* intra-sentencial e intersentencial. Os são-tomenses têm tanto o português como Santome como língua materna. No entanto, o português é a língua mais usada no quotidiano e é alternado com o Santome. A mudança de código entre os são-tomenses, seja ela intra ou intersentencial, se dá primordialmente por uma necessidade de expressão do pensamento que PST, sozinho, não é capaz de fazer. Porém, nem toda suposta alternância intra-sentencial é uma mudança de código. No seu estado atual, muitos itens lexicais do ST fazem parte do PST e são usados naturalmente pelos falantes, que não se dão conta de que são originalmente do ST.

CONSIDERAÇÕES FINAIS

Esta tese teve como objetivo geral apresentar o panorama linguístico de São Tomé e Príncipe, com foco maior no uso cotidiano no Português falado na ilha de São Tomé. A escolha dos temas discutidos surgiu da necessidade do preenchimento de algumas lacunas nos estudos linguísticos relativos ao país ou dos limites alcançados pelos trabalhos realizados por diversos(as) autores(as) até o momento. E o título – Português de São Tomé (e Príncipe): Lexicalização, Gramaticalização e *Code-switching* em contexto multilíngue, foi pensado justamente por entendermos que, no contexto de São Tomé e Príncipe, a pesquisa só ficaria completa e contundente tratando os três tópicos - lexicalização, gramaticalização e *code-switching* – de forma simultânea e/ou paralela, uma vez que eles ocorrem de forma praticamente simultânea e/ou concatenada entre os falantes em razão da coexistência e contato entre as quatro línguas principais – Português, Santome, Ngola e Lung'le.

Para embasamento teórico sobre Lexicalização e Gramaticalização, a pesquisa seguiu os pressupostos de Meillet (2020 [1912]), Moreno Cabrera (1998), Castilho (2003), Neves (2004, 2021), Gonçalves et al. (2007), Bybee (2016, 2020), entre outros, e, línguas em contato e *Code-switching*, foi amparada nas perspectivas de Grosjean (1982), Gumperz (1982), Tarallo & Alkmin (1987), Calvet (2002) e Mozzillo (2009). Esta tese seguiu uma abordagem qualitativa com o propósito de descrever itens lexicais e estruturas sintáticas do Português de São Tomé, assim como o uso do *code-switching* (Português-Santome) nas interações sociais entre os falantes. Para tanto, o corpus utilizado compôs-se de dados coletados em textos multimodais, incluindo artigos de

jornais digitais locais, livros de contos tradicionais e história de São Tomé e Príncipe, uma crônica e, principalmente, publicações em mídias sociais.

A análise morfológica desses itens lexicais, as estruturas sintáticas e os processos de lexicalização e gramaticalização do Português de São Tomé demonstraram que grande parte do léxico e construções sintáticas do Português de São Tomé resultam de empréstimos, decalques e/ou derivações do Santome, motivados pela coexistência e pelo contato linguístico, indicando um impacto importante do Santome no Português de São Tomé. Além disso, o recurso frequente ao *code-switching* Português-Santome por parte dos falantes nessa comunidade de fala é motivado por uma questão de identidade e pertencimento e pela necessidade de expressar pensamentos, necessidades e emoções que o Português de São Tomé, sozinho, não é capaz de realizar na sua plenitude.

Os dados coletados e analisados foram considerados representativos da comunidade de fala do Português de São Tomé (PST), e os resultados, assim como as discussões apresentados nesta tese, demonstram características peculiares do cenário linguístico de São Tomé (e Príncipe). As ocorrências de processos de lexicalização e de gramaticalização observados, sejam originárias de português, sejam decorrentes da utilização do *code-switching* em função do contato linguístico do PST com o ST, revelam a complexidade e as particularidades dos processos de lexicalização e gramaticalização em São Tomé (e Príncipe).

Foram adotados procedimentos metodológicos de natureza qualitativa, para apresentar o contexto sócio-histórico-cultural-linguístico do processo de formação da sociedade são-tomense e analisar os processos de mudança linguística do Português de São Tomé (PST) com foco na Lexicalização, Gramaticalização e no *Code-switching*,

adotando para tanto uma abordagem descritiva de processos e fenômenos linguísticos no PST amparados nos preceitos funcionalistas da Linguística Baseada no Uso.

Os resultados e constatações das análises demonstram que os procedimentos metodológicos adotados foram adequados para os objetivos propostos nesta tese. Além disso, ressaltaram que as mídias sociais representam uma grande e rica fonte de dados passíveis de análise e consideramos que as comunidades digitais são extensões das capacidades físicas e mentais do ser humano, ou seja, as mídias sociais são representativas das comunidades de fala de são-tomenses, principalmente, o Português de São Tomé e o Santomé.

Importa acentuar, no entanto, que a ausência de uma instituição ou órgão permanente que estabeleça as diretrizes de políticas linguísticas e/ou de padronização nos moldes do VOLP (Vocabulário Ortográfico da Língua Portuguesa) vinculado à Academia Brasileira de Letras (ABL), por exemplo, cria uma zona cinzenta principalmente na definição de ocorrências formais de *code-switching*. Em outras palavras, embora boa parte dos ideofones, onomatopeias e ditados populares sejam oriundos do Santomé, argumentamos que são atualmente parte importante do léxico do Português de São Tomé (e Príncipe). Assim sendo, argumentamos que esse é o Português de São Tomé (e Príncipe). “E Príncipe” é aqui reiterado novamente, pois, além das observações já feitas na Introdução desta tese, não é possível delimitar ou distinguir a origem dos usuários (ilha de São Tomé ou ilha do Príncipe) cujas postagens foram utilizadas como base de análise desta pesquisa.

Assim sendo, podemos confirmar a hipótese sustentada nesta tese de que o Português de São Tomé é uma variante da língua portuguesa com características singulares, pois a compreensão e a análise aprofundada dessa língua, assim como os

processos de mudança linguística, incluindo a lexicalização e gramaticalização, dependem, em grande medida, do conhecimento e domínio das línguas nativas do país, principalmente do Santome; o *code-switching* é um recurso de confirmação de identidade, pertencimento e necessidade de expressão de pensamentos e emoções que somente o Santome é capaz de realizar na sua plenitude.

Além disso, este trabalho busca contribuir para a pesquisa linguística de São Tomé e Príncipe em geral, e de São Tomé, em particular, a partir da perspectiva do lugar de fala de um nativo e falante de Português de São Tomé e do Santome. Acreditamos também que essa temática não se esgote com esta tese e que ampliação e aprofundamentos sejam possíveis, podendo esta tese servir de referência para trabalhos futuros.

REFERÊNCIAS

ALBUQUERQUE, Luis de. **A ilha de São Tomé nos séculos XV e XVI**. Alfa, 1989.

ALMEIDA, Nilson Teixeira de. **Gramática completa para concursos e vestibulares**. 2ª ed. – São Paulo: Saraiva, 2009.

ANDRADE, Carolina Queiroz; PACHECO, Cintia da Silva; RODRIGUES, Ulisdete Rodrigues de Souza. **Contato intra e interlinguístico**: variação e mudança linguística nos pronomes pessoais do português brasileiro, português uruguaio e crioulo caboverdiano. In: Kyiria Rebeca Finardi; Maria Marta Pereira Scherre; Leila Maria Tesch; Hebe Macedo de Carvalho. (Org.). *A diversidade de fazeres em torno da linguagem: universidades, faculdades e educação básica em ação*. 1ed. Campinas, SP: Pontes Editores, 2019, v., p. 37-51.

ARAUJO, Gabriel Antunes de. **Ideofones na língua são-tomense**. PAPIA-Revista Brasileira de Estudos do Contato Linguístico, v. 19, n. 1, p. 23-37, 2009.

A SOMBRA do Paço Fiá Gleza o Téla Nón deseja feliz Natal aos seus leitores. Téla Nón, São Tomé, São Tomé e Príncipe, 24 dez 2010. Disponível em: <https://www.telanon.info/cultura/2010/12/24/5939/a-sombra-do-paco-fia-gleza-o-tela-non-deseja-feliz-natal-aos-seus-leitores/>. Acesso em: 10 mar. 2024.

BANCO CENTRAL DE SÃO TOMÉ E PRÍNCIPE. **Nota de cinquenta mil dobras**. São Tomé, 2004.

BARBOSA, Neyara Macedo Coelho. **Parêmsias especializadas em linguagem jurídica bilíngue. (Tese de Doutorado em Linguística)**. Brasília: Instituto de Letras, Universidade de Brasília, 2023, 234 p.

BARBOSA, Rogério Andrade. **Soyas de Sun Tataluga**: histórias que me contaram em São Tomé e Príncipe. 2015.

BASÍLIO, Margarida. **Teoria Lexical**. Editora Ática – São Paulo, 1987.

BOM JESUS, Jorge. **Slogan do Partido MLSTP nas Eleições legislativas de 2022**. Facebook, 12 set. 2022. Disponível em:

<https://www.facebook.com/share/p/zA2yUMm2mXTofbix/?mibextid=oFDknk>. Acesso em: 21 mai. 2024.

BRASÃO de armas de São Tomé e Príncipe. In: WIKIPÉDIA, a enciclopédia livre. Flórida: Wikimedia Foundation, 2021. Disponível em: https://pt.wikipedia.org/w/index.php?title=Bras%C3%A3o_de_armas_de_S%C3%A3o_Tom%C3%A9_e_Pr%C3%ADncipe&oldid=62712776. Acesso em: 21 mai. 2024.

BYBEE, Joan. **Língua, uso e cognição.** Maria Angélica Furtado Cunha. (Trad.). São Paulo: Cortez. CAVALCA, 2016 [2010].

_____. **Mudança linguística.** Tradução Marcos Bagno. Petrópolis: Editora Vozes, 2020 [2015].

CALVET, Louis-Jean. **Sociolinguística: uma introdução crítica.** Tradução: Marcos Marcionilo. São Paulo: Parábola, 2002.

CASTILHO, Ataliba Teixeira de. **Representações das categorias cognitivas e sua diacronia.** Interface Linguística cognitiva - Linguística histórica. *Filologia E Linguística Portuguesa*, 13(1), 63-87, 2011. Disponível em: <https://doi.org/10.11606/issn.2176-9419.v13i1p63-87>, acesso em: 8 abr. 2023.

_____. **Proposta funcionalista de mudança linguística:** Lexicalização, semanticização, discursivização e gramaticalização das preposições do eixo transversal no Português Brasileiro. 2013b, p. 1-32. Disponível em: <https://silo.tips/download/proposta-funcionalista-de-mudana-lingistica>. Acesso em 01 de mar. 2023.

CEN avisa que "casos de banho" serão levados à Justiça. RFI/África Lusófona. São Tomé e Príncipe, 16 de jul. 2021. Disponível em: <https://www.rfi.fr/pt/são-tomé-e-príncipe/20210716-cen-avisa-que-casos-de-banho-serão-levados-à-justiça>. Acesso em: 5 jan. 2023.

COELHO, Margarida. COIA, Cíntia Alves Valentina. LUISELLI, Donata. USELI, Antonella. HAGEMEIJER, Tjerk. AMORIM, António. DESTRO-BISOL, Giovanni. ROCHA, Jorge. **Human microevolution and the Atlantic slave trade:** a case study from Sao Tome. *Current Anthropology*, v. 49, n. 1, p. 134-143, 2008.

COELHO, Sueli. **Expansão gramatical e expansão lexical**: dois processos linguísticos paralelos. In: VITRAL, L. COELHO, S. (org.) Estudos de processos de gramaticalização em português: metodologias e aplicações. Campinas, SP: Mercado de Letras, 2010. p. 333-346.

DAIO, Olinto. **Semplu**: colectânea de provérbios e máximas santomenses. Gesmédia, 2002.

DELANCEY, Scott. **Lectures on functional syntax**. Notes for the Summer School of the Linguistic Society of America held at the University of California at Santa Barbara, 2001.

DÊNDE, Manuel. **Deputados apelam às mulheres a se absterem de usar a expressão “ontá ocê bom”**. Disponível em: <https://www.stp-press.st/2020/11/10/deputados-apelam-as-mulheres-a-se-absterem-de-usar-a-expressao-onta-oce-bom/>. Acesso em: 10 Nov. 2020.

DIK, Simon C. **The theory of functional grammar**. Walter de Gruyter, 1989.

DIRE(C)ÇÃO DOS IMPOSTOS DE SAO TOME E PRINCIPE. **Nova Fatura Vêndê Pleçu Dóxi**. Disponível em: <https://impostos.financas.gov.st/>. Acesso em 06 jun. 2023.

DIRE(C)ÇÃO NACIONAL DA CULTURA DE SAO TOME E PRINCIPE. **Contos tradicionais santomenses**. São Tomé: DNC, 1984. 79 p.

FAULSTICH, Enilde. **Efeitos da (nova) ortografia no léxico do português**: mecanismos gramaticais na grafia de algumas palavras e resultados no uso. In LOBO, T., CARNEIRO, Z., SOLEDADE, J., ALMEIDA, A., and RIBEIRO, S., orgs. *Rosae*: linguística histórica, história das línguas e outras histórias [online]. Salvador: EDUFBA, 2012, pp. 363-380. ISBN 978-85-232-1230-8. Disponível em: <https://books.scielo.org/id/67y3k/pdf/lobo-9788523212308-26.pdf>>. Acesso em 12 out. 2025.

FORTUNATO, Isabella Venceslau. **Gramaticalização e lexicalização das lexias complexas no português arcaico**. In: José Sueli MAGALHÃES; Luiz Carlos TRAVAGLIA. (Org.). *Múltiplas Perspectivas em Linguística*. 01 ed. Uberlândia: Edufu, 2008, p. 1394-1403.

KUA NON. **Fulumento**. Facebook, 29 mai. 2025. Disponível em: <https://www.facebook.com/photo.php?fbid=1278039844328905&set=pb.100063687037480.-2207520000&type=3>. Acesso em: 29 set. 2025.

FACEBOOK (2023). **Postagem de JB.** Disponível em: <https://www.facebook.com/joimilte.bonfim>. Acesso em: 29 mai. 2025.

GARRIDO, Hilário. **“Debatofobia”/Uêchaismo.** Tela Non, São Tomé, São Tomé e Príncipe, 07 ago. 2008. Análise. Disponível em: <https://www.telanon.info/suplemento/analise/2008/08/07/285/debatofobia/>. Acesso em: 5 jan. 2023.

GONÇALVES, Rita; HAGEMEIJER, Tjerk. **O português num contexto multilíngue: o caso de São Tomé e Príncipe.** Lisboa, 2015.

GONÇALVES, Sebastião Carlos Leite; LIMA-HERNANDES, Maria Célia; CASSEB-GALVÃO, Vânia Cristina (org.). **Introdução à gramaticalização: princípios teóricos e aplicação.** São Paulo: Parábola Editorial, 2007. 207p. ISBN: 978-85-88456-70x.

GRAÇA, Ramusel. **Forte de São Sebastião/Estátuas dos navegadores portugueses João de Santarém, Pêro Escobar e João de Paiva.** Disponível em: <<https://www.dw.com/pt-002/são-sebastião-a-fortaleza-que-counta-a-história-de-são-tomé/g-46225031>>. Acesso em: 15 mai. 2024.

GRANDA, Germán de. 1970. **Un temprano testimonio sobre las hablas 'criollas' en África y América (P. Alonso de Sandoval, De instauranda Aethiopum salute, Sevilla, 1627).** Thesaurus 25: 1-II, 1970.

GROSJEAN, François. **Life with two languages: an Introduction to Bilingualism.** Harvard University Press, 1982.

GUMPERZ, John Joseph. **Discourse Strategies.** Cambridge: Cambridge University Press, 1982.

HAGEMEIJER, Tjerk. 2009. **As línguas de S. Tomé e Príncipe.** Revista de Crioulos de Base Lexical Portuguesa e Espanhola, Vol. 1.1 (27 pgs). Disponível em: https://cvc.cervantes.es/lengua/thesaurus/pdf/25/TH_25_001_001_1.pdf. Acesso em 10 abr. 2023.

HAGEMEIJER, Tjerk. AMORIM, António. DESTRO-BISOL, Giovanni. ROCHA, Jorge. **“Human Microevolution and the Atlantic Slave Trade: A Case Study from São Tomé.”**

Current Anthropology, vol. 49, no. 1, 2008, pp. 134–43. JSTOR, <https://doi.org/10.1086/524762>. Acesso em 19 Nov. 2023.

HAGEMEIJER, Tjerk; Hendrickx, Iris; Amaro, Haldane; Tiny, Abigail. 2012. **A Corpus of Santome**. In Proceedings of the SALTML-AfLaT workshop, Istanbul, Turkey, 2012. European Language Resources Association (ELRA), 61-66.

HEINE, Bernd; CLAUDI, Ulrike; HÜNNEMEYER, Friederike. **Grammaticalization: a conceptual framework**. Chicago: the University of Chicago Press, 1991b. Disponível em: http://www.prof-bernd-heine.de/books/Heine_etal_1991_Grammaticalization.pdf. Acesso em: 13 out. 2025.

HOPPER, Paul J; TRAUGOTT, Elizabeth Closs. **Globalization**. Cambridge: Cambridge University Press, 1993.

HOTEL Hunter (2024). **Hotel Omali São Tomé**. Disponível em: <https://hotelhunter.com/en/sao-tome-and-principe/sao-tome/omali-sao-tome>. Acesso em 21 mai. 2024.

INSTITUTO NACIONAL DE ESTATÍSTICA. **Características Educacionais de População, RGPH-2001**. São Tomé: INE, 2003a. 81 p.

_____. **Estado e Estrutura da População em S. Tomé e Príncipe**, RGPH-2012. São Tomé: INE, 2014. 68 p.

_____. **Migrações**, RGPH-2001. - São Tomé: INE, 2003b, 45 p.

KARASCH, Mary. **A vida dos escravos no Rio de Janeiro (1808-1850)**. São Paulo: Companhia das Letras, 2000.

LIMA, Pepê. **Sun Dentxi Bétu**. YouTube, 3 set. 2012. Disponível em: <https://www.youtube.com/watch?v=yze8NYD-qrY>. Acesso em: 18 set. 2025.

LIMA, Pepê. **Dentxi Bétu 2**. YouTube, 20 nov. 2018. Disponível em: <https://www.youtube.com/watch?v=2el4MrA2aGw>. Acesso em: 18 set. 2025.

MACARA, Teresa. LISBOA, Luísa (orgs.). **Sóias e Contagi**: quem conta um conto...aumenta um ponto! A língua portuguesa nos contos tradicionais santomenses. Ed. AMU/Flimá – Lisboa, 2009.

MARTELOTTA, Mário Eduardo. **Mudança linguística**: uma abordagem baseada no uso. São Paulo: Cortez Editora, 2011.

MARTELOTTA, Mário Eduardo; VOTRE, Sebastião Josué; CEZARIO, Maria Maura. **Gramaticalização no português do Brasil**. Rio de Janeiro: Tempo Brasileiro, 1996.

MARTINET, André. **Que c'est la linguistique fonctionnelle?** ALFA, v. 38, pp. 11-18, 1994.

MEILLET, Antoine. **L'évolution des formes grammaticales**. Scientia (Rivista di Scienza), v. 12, n. 26), p. 6, 1912.

MINISTÉRIO DA SAÚDE. **Estratégia integrada de Saúde Reprodutiva, Materna, Neonatal, Infantil e do Adolescente e Nutrição 2019-2023**. São Tomé, 2018. 77 p. Disponível em: https://saotomeandprincipe.unfpa.org/sites/default/files/pub-pdf/plano_sr_2019_2023.pdf. Acesso em: 15 out. 2025.

MORENO CABRERA, Juan Carlos. **On the relationship between grammaticalization and lexicalization**. In: The limits of grammaticalization. John Benjamins Publishing Company, 2011. p. 211-227.

MOZZILLO, Isabella. **O code-switching: fenômeno inerente ao falante bilíngue**. PAPIA: Revista Brasileira de Estudos Crioulos e Similares, São Paulo, v. 19, p. 85-200, 2009.

NEVES, Maria Helena de Moura. **Gramática funcional**: interação, discurso e texto. São Paulo: Contexto, 2021.

NEVES, Maria Helena de Moura et al. **Uma introdução ao funcionalismo**: proposições, escolas, temas e rumos. Funcionalismo e gramaticalização: teoria, análise e ensino. João Pessoa: Ideia, p. 13-28, 2004.

NILEP, Chad. **“Code Switching” in Sociocultural Linguistics**. Colorado Research in Linguistics, [S. l.], v. 19, 2006. DOI: 10.25810/hnq4-jv62. Disponível em: <https://journals.colorado.edu/index.php/cril/article/view/273>. Acesso em: 22 nov. 2023.

PIMENTEL NETO, Julwaity. **Ba kwê**. Messenger, 22 dez. 2024. Disponível em: <https://www.facebook.com/messages/e2ee/t/26603258789287663>. Acesso em: 01 jul. 2025.

REIS, Fernando. **Histórias da roça**. Sociedade de Expansão Cultural, 1970.

SANTIAGO, Ana Maria. AGOSTINHO, Ana Livia. (2020). **Situação linguística do português em São Tomé e Príncipe**. *A Cor Das Letras*, 21(1), 39–61. Disponível em: <https://doi.org/10.13102/cl.v21i1.4970>. Acesso em 28 fev. 2024.

SÃO TOMÉ E PRÍNCIPE. **Constituição (2003)**. Constituição da República Democrática de São Tomé e Príncipe. São Tomé: Assembleia Nacional, 2003.

_____. **Decreto nº 19, de 14 de agosto de 2013**. Dispõe sobre a ortografia e cria o Alfabeto Unificado para as Línguas Nativas de São Tomé e Príncipe, designadamente o Santome, Ngola e o Lung'le. São Tomé, 2013.

_____. Assembleia Nacional. **O regime democrático em S. Tomé e Príncipe e seus antecedentes históricos**. Disponível em: <http://www.parlamento.st/>. Acesso em: 20 mai. 2023.

SÃO TOMÉ (SÃO TOMÉ E PRÍNCIPE). In: WIKIPÉDIA, a enciclopédia livre. Flórida: Wikimedia Foundation, 2023. Disponível em: [https://pt.wikipedia.org/w/index.php?title=S%C3%A3o_Tom%C3%A9_\(S%C3%A3o_Tom%C3%A9_e_Pr%C3%ADncipe\)&oldid=67140805](https://pt.wikipedia.org/w/index.php?title=S%C3%A3o_Tom%C3%A9_(S%C3%A3o_Tom%C3%A9_e_Pr%C3%ADncipe)&oldid=67140805). Acesso em: 19 dez. 2023.

SEIBERT, Gerhard. **A questão da origem dos Angolares de São Tomé**. Instituto Superior de Economia e Gestão – CEsA Brief papers nº 5-1998. Disponível em: <https://www.repository.utl.pt/handle/10400.5/2112>. Acesso em 18 nov. 2023.

_____. **Colonialismo em São Tomé e Príncipe**: hierarquização, classificação e segregação da vida social. *Anuário Antropológico*, v. 40, n. 2, p. 99-120, 2015.

_____. **Os angolares da Ilha de São Tomé: náufragos, autóctones ou quilombolas?**. Textos de história, v. 12, n. 1-2, p. 43-64, 2004.

SILVA, Lourenço da. **Presidente inaugura estátua do Rei Amador, a figura emblemática da história são-tomense (2018)**. Disponível em: <https://www.stp-press.st/2018/11/02/presidente-inaugura-estatua-do-rei-amador-figura-emblematica-da-historia-sao-tomense/>. Acesso em: 15 mai. 2024.

SILVA, Lourenço da. **Vava Neviense, Melhor Cantor do Ano e Rádio Nacional volta ser a Melhor Instituição do Ano 2021**. Disponível em: <https://www.stp-press.st/2021/01/02/vava-neviense-melhor-cantor-do-ano-e-radio-nacional-volta-ser-a-melhor-instituicao-do-ano/>. Acesso em: 21 ago. 2025.

SOUZA, Alan de Carvalho. **A importância do café para São Tomé e Príncipe frente à proibição do comércio de escravizados pela Inglaterra**. Afro-Ásia, Salvador, n. 63, 2021. DOI: 10.9771/aa.v0i63.38370. Disponível em: <https://periodicos.ufba.br/index.php/afroasia>. Acesso em: 17 nov. 2023.

TAVARES, Miguel Sousa. **Equador**. Editora Companhia das Letras, 2004.

VIAGEM a São Tomé (2023). **Vianteiro**. Disponível em: <https://www.facebook.com/viagemasaotome/posts/extraindo-o-fabuloso-delicioso-vinho-da-palmasaotome-beach-tours-viagemeturismo-/736450475171742/>. Acesso em: 21 set. 2025.

WORLDOMETERS (2023). **Mapa político de STP**. Disponível em: <https://www.worldometers.info/maps/sao-tome-and-principe-political-map/>. Acesso em: 22 mai. 2024.