



UFPE

UNIVERSIDADE FEDERAL DE PERNAMBUCO

CENTRO DE ARTES E COMUNICAÇÃO

DEPARTAMENTO DE LETRAS

PROGRAMA DE PÓS-GRADUAÇÃO EM LETRAS

GABRYELLA FRAGA DE OLIVEIRA

**DISTRIBUIÇÃO DE SUJEITO NULO E PLENO NO PORTUGUÊS FALADO E
ESCRITO POR UNIVERSITÁRIOS GUINEENSES E SÃO-TOMENSES DA
UNILAB-CEARÁ: interferências do português brasileiro?**

Recife

2019

GABRYELLA FRAGA DE OLIVEIRA

**DISTRIBUIÇÃO DE SUJEITO NULO E PLENO NO PORTUGUÊS FALADO E
ESCRITO POR UNIVERSITÁRIOS GUINEENSES E SÃO-TOMENSES DA
UNILAB-CEARÁ: interferências do português brasileiro?**

Dissertação apresentada ao Programa de Pós-Graduação em Letras da Universidade Federal de Pernambuco, como requisito parcial para a obtenção do título de Mestra em Letras.

Área de concentração: Linguística

Orientadora: Prof^a. Dr^a. Cláudia Roberta Tavares Silva

Recife

2019

O48d	Oliveira, Gabryella Fraga de Distribuição de sujeito nulo e pleno no português falado e escrito por universitários guineenses e são-tomenses da UNILAB-Ceará: interferências do português brasileiro? / Gabryella Fraga de Oliveira. – Recife, 2019. 142f.: il. Orientadora: Cláudia Roberta Tavares Silva. Dissertação (Mestrado) – Universidade Federal de Pernambuco. Centro de Artes e Comunicação. Programa de Pós-Graduação em Letras, 2019. Inclui referências. 1. Parâmetro do sujeito nulo. 2. Contato linguístico. 3. Português são-tomense. 4. Português guineense. I. Silva, Cláudia Roberta Tavares (Orientadora). II. Título.
410	CDD (22. ed.) UFPE (CAC 2019-253)

GABRYELLA FRAGA DE OLIVEIRA

**DISTRIBUIÇÃO DE SUJEITO NULO E PLENO NO PORTUGUÊS FALADO E
ESCRITO POR UNIVERSITÁRIOS GUINEENSES E SÃO-TOMENSES DA
UNILAB-CEARÁ: interferências do português brasileiro?**

Dissertação apresentada ao Programa de Pós-Graduação em Letras da Universidade Federal de Pernambuco, como requisito parcial para a obtenção do título de Mestra em Letras.

Aprovada em: 18/02/2019.

BANCA EXAMINADORA

Prof^a. Dr^a. Claudia Roberta Tavares Silva (Orientadora)
Universidade Federal de Pernambuco

Prof^a. Dr^a. Silvana Silva de Farias Araújo (Examinadora Externa)
Universidade Estadual de Feira de Santana

Prof.^o Dr. Adeilson Pinheiro Sedrins (Examinador Externo)
Universidade Federal Rural de Pernambuco

Dedico esta dissertação a mainha e a painho, tudo sempre foi por vocês, ao meu marido e companheiro, Lucas, você foi essencial, à minha orientadora, Cláudia Roberta, às minhas amigas Elizabeth e Thayse, e ao meu cãopanheiro, Dexter, com todo carinho e gratidão do mundo.

AGRADECIMENTOS

A Deus, por toda força e sabedoria concedida nos momentos que pensei que desistiria, tudo que foi preciso para concluir essa dissertação vem d'Ele.

Aos meus pais, Joelma e Jailson, que me apoiaram incondicionalmente durante meu trajeto e me incentivaram sempre a estudar, em especial, à minha mãe que sempre disse que “o conhecimento é a única coisa que ninguém nunca vai poder tirar de mim”.

A Lucas, meu amigo e parceiro de vida, por ter encarado essa empreitada ao meu lado em todas as horas e por toda compreensão, apoio e carinho nos momentos difíceis.

À minha orientadora, Cláudia Roberta, por toda dedicação, compreensão, incentivo, palavras de consolo, por ter acreditado na minha capacidade e pelo pensamento sempre positivo de que tudo ia dar certo. A senhora é especial demais!

À profa. Silvana Araújo, por acompanhar este trabalho desde o período da qualificação e pelas palavras incentivadoras.

À Profa. Silvana Araújo e ao professor Adeilson Sedrins, por compor a banca avaliadora deste trabalho.

Às professoras Stella Virgínia Telles, da Universidade Federal de Pernambuco (UFPE), e Telma Moreira Viana Magalhães, da Universidade Federal de Alagoas (UFAL), por terem aceitado fazer parte da banca, como membros suplentes.

A Elizabhatt Costa e Thayse Paraíso, pelas horas de alegria e desespero compartilhadas, traçar esse caminho não foi fácil, mas foi muito mais doce com vocês ao meu lado.

A Déreck, Juliana, Edrielly e Jamilys, amigos que encontrei na pós e que trilharam esse caminho comigo, compartilhando dúvidas, vitórias, sorrisos, sonhos e conhecimento. Levo vocês para a minha vida.

Às minhas amigas, Manuela, Nauane, Laís, Amanda e Patrícia que torceram por esse título desde quando eu disse que faria a prova de seleção, por todos pensamentos positivos e pela compreensão da minha ausência em vários momentos em que não pude estar, pois estava ocupada com a dissertação.

À minha sogra Laudiceia e ao meu sogro Valdomiro, por todas as orações e votos de que tudo ia dar certo.

A Dexter, por me acompanhar em todas as noites de estudo e escrita, esse processo não foi solitário, ele esteve ao meu lado, dando carinho e companhia sempre.

Às minhas tias, Marcia, Adriana, Flávia e Kátia, por todo apoio, incentivo, carinho e orgulho que sempre sentiram por mim.

Aos professores que ministraram às disciplinas cursadas na pós, Sivaldo, Virgínia, Xavier, Joyce, Kazuê, Masip, por todo conhecimento compartilhado.

Aos informantes que me receberam muito bem na UNILAB-CE, contribuíram imensamente e foram fundamentais para essa pesquisa.

À professora Artemisa Odila, por me ensinar tanto sobre a vida em tão pouco tempo.

A todos os funcionários que compõem o Programa de Pós-Graduação em Letras da Universidade Federal de Pernambuco.

A CAPES, pelo apoio financeiro a esta pesquisa.

Eu acho que a língua portuguesa é hoje, talvez, uma das línguas europeias com maior vivacidade, com maior dinamismo. Não por causa de nenhuma essência especial do Português, mas por causa de uma razão histórica que aconteceu no Brasil, em que Portugal deu origem a um filho maior que o próprio pai. (COUTO, 2010)

RESUMO

No continente africano, o português convive com várias línguas africanas, resultando assim em um intenso contexto de multilinguismo. Recentes investigações linguísticas apontam para o fato de que a norma urbana culta nos países africanos é muito próxima à norma padrão do português europeu (PE) (cf. MOTA; MIGUEL; MENDES, 2012; BRANDÃO; VIEIRA, 2012; GONÇALVES, 2000). Além disso, a maior ou menor utilização de outra língua favorece, respectivamente, uma maior ou menor proximidade com a variedade europeia do português, língua oficial em alguns países africanos. Diante disso e baseando-nos no modelo teórico de Princípios e Parâmetros (CHOMSKY, 1981, 1986), centramos nossa atenção neste trabalho na distribuição de sujeitos nulos e plenos em dados orais e escritos produzidos por falantes guineenses e são-tomenses que moram no Brasil e estudam na Universidade de Integração Internacional da Lusofonia Afro-Brasileira (UNILAB), localizada na cidade de Redenção-CE, comparando nossos resultados com os já obtidos no PE, em que sujeitos nulos são muito produtivos e analisando as possíveis interferências do português brasileiro (PB). Para tanto, os dados que compõem o corpus (contextos declarativos finitos) foram selecionados de entrevistas informais e textos narrativos produzidos por vinte e quatro alunos guineenses e vinte e quatro alunos são-tomenses, distribuídos entre homens e mulheres. As variáveis selecionadas para a análise são: país de origem, sexo, período do curso (tempo de permanência no Brasil), a pessoa gramatical, tipo de oração, o traço do referente de terceira pessoa e a duplicação do sujeito. Os resultados apontam que há uma maior incidência de sujeitos nulos, embora verifiquemos que há pontos de divergência entre os resultados dos dois países em questão. Ao que percebemos, essa divergência, pode ocorrer por questões relacionados ao contato linguístico: em Guiné Bissau, o português é uma L2, ao passo que em São Tomé é muito frequentemente uma L1 (cf. HAGEMEIJER 2009). Também percebemos que o português guineense, em sua modalidade falada, tem preferência por sujeitos plenos em contextos semelhantes aos que o PB faz essa escolha, enquanto que o português são-tomense é uma língua que vai ao encontro da variedade europeia do português e é favorável à produção de sujeitos nulos, mesmo em contextos em que o PB teria preferência por sujeitos plenos.

Palavras-chave: Parâmetro do sujeito nulo. Contato linguístico. Português são-tomense. Português guineense.

RESUMEN

En el continente africano, el portugués convive con lenguas africanas variadas, como resultado se estableció un intenso contexto de multilingüismo. Investigaciones lingüísticas recientes señalan el hecho de que la norma culta urbana en los países africanos es muy cercana a la norma patrón del portugués europeo (PE) (cf. MOTA; MIGUEL; MENDES, 2012; BRANDÃO; VIEIRA, 2012; GONÇALVES, 2000). Además de eso, la mayor o menor utilización de otra lengua, favorece, respectivamente, una mayor o menor proximidad con la variedad europea del portugués, lengua oficial en algunos países africanos. Frente a eso y adoptando por base teórica el modelo de Principios y Parámetros (CHOMSKY, 1981, 1986), centramos nuestra atención, en este trabajo, en la distribución de sujetos nulos y plenos en datos orales y escritos producidos por hablantes bisauguineanos y santotomenses que viven en Brasil y estudian en la Universidade de Integração Internacional da Lusofonia Afro-Brasileira (UNILAB), ubicada en la ciudad de Redenção-CE, comparando nuestros resultados con aquellos ya obtenidos en el PE, en que sujetos nulos son muy productivos y analizando las posibles interferencias del Portugués Brasileño (PB). En consecuencia, los datos que componen el corpus (contextos declarativos finitos) fueron seleccionados de entrevistas informales y textos narrativos producidos por veinticuatro alumnos bisauguineanos y veinticuatro alumnos santotomenses, distribuidos entre hombres y mujeres. Las variables seleccionadas para el análisis son: país de nacimiento, sexo, período de curso (tiempo de permanencia en Brasil), la persona gramatical, tipo de oración, el rasgo del referente de tercera persona y la duplicación del sujeto. Los resultados señalan que hay una mayor incidencia de sujetos nulos, mismo que verifiquemos que existen puntos de divergencia entre los resultados de los dos países presentados. Por lo que pudimos percibir, esa divergencia, puede ocurrir por cuestiones relacionadas al contacto lingüístico: en Guinea-Bisáu, el portugués es una L2, mientras que en Santo Tomé es frecuentemente una L1 (cf. HAGEMEIJER 2009). Se debe agregar que el portugués bisauguineano, en su modalidad oral, posee preferencia de sujetos plenos en contextos semejantes al PB. Por otro lado, el portugués santotomense es una lengua que se opone a la variedad europea del portugués y es favorable a la producción de sujetos nulos, incluso en las situaciones en que el PB manifestaría preferencia por sujetos plenos.

Palabras-clave: Parámetro del sujeto nulo. Contato lingüístico. Portugués santotomense. Portugués bisauguineano.

LISTA DE GRÁFICOS

Gráfico 1 - Ocorrência de sujeito nulo em PB através de sete períodos. (%).....	39
Gráfico 2 - Ocorrência de sujeitos nulos nas três pessoas em PB. (%).....	39
Gráfico 3 - Quantitativo geral e percentual de sujeitos nulos e plenos nas variedades guineense e são-tomense do português.....	84
Gráfico 4 - Ocorrência de sujeitos nulos e plenos nas variedades do português de Guiné Bissau e São Tomé e Príncipe (%).....	85
Gráfico 5 - Percentual de sujeitos nulos e plenos, realizados no PGB, tomando por base a variável sexo.....	86
Gráfico 6 - Percentual de sujeitos nulos e plenos, realizados na variedade são-tomense do português, tomando por base a variável sexo.....	88
Gráfico 7 - Percentual de sujeitos nulos e plenos, realizados na variedade guineense do português, tomando por base a variável período do curso.....	89
Gráfico 8 - Percentual de sujeitos nulos e plenos, realizados na variedade são-tomense do português, tomando por base a variável período do curso.....	90
Gráfico 9 - Ocorrência de sujeitos nulos segundo o tipo de oração (%).....	96
Gráfico 10 - Ocorrência de sujeito nulo com base nas orações: independentes, principais e relativas (%).....	97
Gráfico 11 - As ocorrências de sujeito nulo com base no traço semântico do referente (%).....	99
Gráfico 12 - Ocorrência de sujeitos nulos no PST, segundo o tipo de oração (%).....	104
Gráfico 13 - Ocorrência de sujeito nulo com base nas orações: independentes, principais e relativas (%).....	105
Gráfico 14 - As ocorrências de sujeito nulo com base no traço semântico do referente (%).....	106
Gráfico 15 - Quantitativo geral e percentual de sujeitos nulos e plenos, nos dados de fala, do PGB e do PST.....	108
Gráfico 16 - Ocorrência de sujeitos nulos e plenos no PDG e no PST, tomando por base os dados de fala (%).....	109

Gráfico 17 - Percentual de sujeitos nulos e plenos, realizados no PGB, tomando por base a variável sexo.....	110
Gráfico 18 - Percentual de sujeitos nulos e plenos, realizados no PST, tomando por base a variável sexo.....	112
Gráfico 19 - Percentual de sujeitos nulos e plenos, realizados no PGB, tomando por base a variável período do curso.....	113
Gráfico 20 - Percentual de sujeitos nulos e plenos, realizados no PST, tomando por base a variável período do curso.....	114
Gráfico 21 - Ocorrência de sujeito nulo, na língua falada do PGB, segundo a pessoa gramatical e o período do curso (%).....	116
Gráfico 22 - Uso da expressão pronominal A gente e do pronome Nós plenos nos primeiros períodos do PGB.....	118
Gráfico 23 - Uso da expressão pronominal A gente e do pronome Nós plenos, nos últimos períodos do PGB.....	118
Gráfico 24 - Ocorrência de sujeitos nulos nos dados de fala do PGB, segundo o tipo de oração (%).....	122
Gráfico 25 - Ocorrência de sujeito nulo em dados de fala do PGB com base nas orações: independentes, principais e relativas (%).....	123
Gráfico 26 - As ocorrências de sujeito nulo com base no traço semântico do referente (%).....	124
Gráfico 27 - ocorrência de sujeitos nulos segundo o tipo de oração (%).....	129
Gráfico 28 - Ocorrência de sujeito nulo com base nas orações: independentes, principais e relativas (%).....	129
Gráfico 29 - As ocorrências de sujeito nulo com base no traço semântico do referente (%).....	131

LISTA DE QUADROS

Quadro 1 - Língua(s) mais usada(s) e língua oficial em cinco países africanos. Extraído de Oliveira e Silva 2018.....	22
Quadro 2 - Sistema de sujeitos nulos com base nos estudos de Holmberg, Nayudu e Sheehan (2009) e Roberts e Holmberg (2010).....	26
Quadro 3 - Paradigma de flexão verbal extraído de Galves (2001, p. 103).....	35
Quadro 4 - Ocorrência de sujeito nulo em PE, adaptado de Duarte (1995, p. 8).....	36
Quadro 5 - Paradigma de flexão verbal do português brasileiro, adaptado de Duarte (1995, p. 32).....	37
Quadro 6 - Paradigma de flexão verbal do português brasileiro, adaptado de Duarte (1995, p. 32).....	38
Quadro 7 - Paradigma de flexão verbal do português brasileiro, adaptado de Duarte (1995, p. 32).....	38
Quadro 8 - Número de ocorrências das formas pronominais na posição de sujeito em PA em 369 sentenças. Adaptado de Oliveira e Santos (2007, p. 97).....	43
Quadro 9 - Paradigma flexional do kimbundo, adaptado de Teixeira (2012, p. 10).....	44
Quadro 10 - Ocorrência de sujeitos nulos em PM, PB e PE extraída de Bravin dos Santos (2010, p. 31).....	46
Quadro 11 - Distribuição dos dados de concordância (marcação) e não concordância (não marcação) nominal de número e verbal de 3ª pessoa plural nas três variedades estudadas. Adaptado de Brandão e Vieira (2012, p. 1040).....	50
Quadro 12 - Quantitativo geral e percentual de sujeitos plenos e nulos. Extraído de Oliveira e Silva (2018).....	51
Quadro 13 - Paradigma de flexão verbal do crioulo guineense. Adaptado de Castro (2013, p. 102).....	51
Quadro 14 - Paradigmas de flexão verbal do italiano. Elaboração própria.....	60
Quadro 15 - Paradigmas de flexão verbal do inglês. Elaboração própria.....	61
Quadro 16 - Relação dos informantes da pesquisa de acordo com curso, sexo e variedade africana do português.....	75

Quadro 17 - Quantitativo de dados obtidos de acordo com a modalidade de língua e a variedade africana do português.....	79
Quadro 18 - Ocorrência de sujeito nulo, na modalidade escrita da língua, na variedade guineense, segundo a pessoa do discurso.....	92
Quadro 19 - Ocorrência de sujeito nulo, na modalidade escrita do PST, segundo a pessoa gramatical.....	102
Quadro 20 - Ocorrência de sujeitos nulos de primeira pessoa no PGB, tomando por base os primeiros e últimos períodos (%).....	117
Quadro 21 - Ocorrência de sujeitos nulos de segunda pessoa no PGB(%).....	119
Quadro 22 - Ocorrência de sujeitos nulos de terceira pessoa nos dados de fala do PGB (%).....	121
Quadro 23 - Ocorrência de sujeito nulo, na modalidade falada da língua, na variedade são-tomense, segundo a pessoa gramatical (%).....	127

LISTA DE SIGLAS

PA	Português angolano
PCV	Português cabo-verdiano
PGB	Português guineense
PM	Português moçambicano

SUMÁRIO

1	INTRODUÇÃO.....	18
2	O PANORAMA DO ESTUDO.....	22
2.1	DELIMITAÇÃO DO OBJETO DE ESTUDO.....	22
2.2	O CONTATO LINGUÍSTICO E AS VARIEDADES AFRICANAS DO PORTUGUÊS.....	29
2.2.1	A situação sociolinguística de São Tomé e Príncipe.....	31
2.2.2	A situação sociolinguística de Guiné Bissau.....	33
2.3	SOBRE O PARÂMETRO DO SUJEITO NULO EM VARIEDADES DO PORTUGUÊS.....	35
2.3.1	O parâmetro do sujeito nulo no português europeu.....	35
2.3.2	O parâmetro do sujeito nulo no português brasileiro.....	36
2.3.2.1	A proposta de Duarte (1993, 1995).....	37
2.3.2.2	A proposta de Rodrigues (2004), Holmberg (2005) e Kato e Duarte (2014).....	41
2.4	O PARÂMETRO DO SUJEITO NULO E A MORFOLOGIA DE FLEXÃO VERBAL NAS VARIEDADES AFRICANAS DO PORTUGUÊS.....	42
2.4.1	Angola.....	43
2.4.2	Moçambique.....	45
2.4.3	Cabo Verde.....	47
2.4.4	São Tomé e Príncipe.....	49
2.4.5	Guiné Bissau.....	51
2.5	OBJETIVOS.....	52
2.6	HIPÓTESES.....	53
3	FUNDAMENTAÇÃO TEÓRICA.....	54
3.1	PRESSUPOSTOS BÁSICOS DA TEORIA GERATIVA.....	54
3.1.1	Fundamento filosófico.....	54
3.1.2	Definição de linguagem.....	55
3.1.3	Sobre a Faculdade da Linguagem.....	56
3.1.4	Sobre a Gramática Universal e o Modelo de Princípios e Parâmetros.....	58
3.1.4.1	O Parâmetro do Sujeito Nulo.....	60
3.2	AQUISIÇÃO DE L1 E L2 À LUZ DA TEORIA GERATIVA.....	64
3.2.1	Aquisição de L1 à luz da Teoria Gerativa.....	64

3.2.1.1	Aquisição de português como L1 em São Tomé e Príncipe.....	65
3.2.2	Aquisição de L2 à luz da teoria gerativa.....	67
3.2.2.1	Aquisição da linguagem à luz da crioulistica.....	69
3.2.2.2	Aquisição de português como L2 em Guiné Bissau.....	71
4	METODOLOGIA.....	74
4.1	MÉTODOS DE ABORDAGEM E PROCEDIMENTO.....	74
4.2	PERFIL DOS INFORMANTES.....	74
4.3	COLETA DE DADOS.....	76
4.4	CONSTITUIÇÃO DO <i>CORPORA</i>	78
4.5	SELEÇÃO E CODIFICAÇÃO DAS VARIÁVEIS.....	79
4.6	TRATAMENTO QUANTITATIVO DOS DADOS.....	83
5	ANÁLISE DOS DADOS DE ESCRITA.....	84
5.1	COMPARANDO GUINÉ BISSAU E SÃO TOMÉ E PRÍNCIPE: QUANTITATIVO GERAL DOS DADOS.....	84
5.2	AS VARIÁVEIS LINGUÍSTICAS: GUINÉ BISSAU.....	91
5.2.1	Pessoa gramatical e desinência verbal.....	91
5.2.2	A estrutura da oração e seu estatuto sintático dentro do período.....	95
5.2.3	O traço do referente de terceira pessoa.....	99
5.2.4	A duplicação do sujeito.....	100
5.3	AS VARIÁVEIS LINGUÍSTICAS: SÃO TOMÉ E PRÍNCIPE.....	102
5.3.1	Pessoa gramatical e desinência verbal.....	102
5.3.2	A estrutura da oração e seu estatuto sintático dentro do período.....	104
5.3.3	O traço do referente de terceira pessoa.....	106
6	ANÁLISE DOS DADOS DE FALA.....	108
6.1	COMPARANDO GUINÉ BISSAU E SÃO TOMÉ E PRÍNCIPE: QUANTITATIVO GERAL DOS DADOS.....	108
6.2	AS VARIÁVEIS LINGUÍSTICAS: GUINÉ BISSAU.....	115
6.2.1	Pessoa gramatical e desinência verbal.....	115
6.2.2	A estrutura da oração e seu estatuto sintático dentro do período.....	121
6.2.3	O traço referente de terceira pessoa.....	124
6.2.4	A duplicidade do sujeito.....	126
6.3	AS VARIÁVEIS LINGUÍSTICAS: SÃO TOMÉ E PRÍNCIPE.....	126
6.3.1	Pessoa gramatical.....	126
6.3.2	A estrutura da oração e seu estatuto sintático dentro do período.....	128

6.3.3	O traço do referente de terceira pessoa.....	130
7	CONSIDERAÇÕES FINAIS.....	132
	REFERÊNCIAS.....	136

1 INTRODUÇÃO

Em alguns países da África o português é secundarizado e convive com diversas outras línguas, mesmo gozando do status de língua oficial nestes países. São Tomé e Príncipe e Guiné Bissau são dois países que exemplificam diferentes pontos desse complexo multilinguismo, pois, enquanto o português é falado de forma mais frequente nas ilhas de São Tomé e Príncipe, a língua que predomina em Guiné Bissau é o crioulo guineense. Segundo Hagemijer (2009, p. 19-20), São Tomé e Príncipe “é o “único país da África de língua portuguesa onde a maioria da população tem actualmente o Português como primeira língua, havendo assim condição para a emergência de uma nova variedade”.

Segundo Duarte (2006), somente no século XIX é que se instaura o ensino de língua portuguesa em Guiné Bissau, o que acaba conferindo ao crioulo guineense uma estrutura mais estabilizada em relação ao português e o faz ser mais frequentemente utilizado nos mais diversos contextos comunicativos. O nível de escolaridade do falante também é um fator determinante nesse contexto de multilinguismo. Segundo a análise de Brandão e Vieira (2012), no que diz respeito ao domínio da concordância verbal e nominal na fala de são-tomenses, os falantes com nível superior de escolaridade e que fazem menos uso da variedade do crioulo de São Tomé (o santomé), apresentam uma maior tendência a reproduzir uma concordância mais afinada com o português europeu.

Tomando como base o exposto até aqui, temos que o foco deste trabalho atenta para a comparação entre dados de fala e escrita dos falantes guineenses e são-tomenses que residem atualmente no Brasil em função de seus cursos de graduação na Universidade de Interação Internacional de Lusofonia Afro Brasileira (UNILAB), situada no estado do Ceará, mais especificamente na cidade de Redenção. Levamos em conta a distribuição de sujeitos nulos e plenos em dados orais e escritos produzidos por estes alunos, observando também possíveis interferências decorrentes do contato com o português brasileiro. Como base para nossa análise, usaremos o Modelo de Princípios e Parâmetros (CHOMSKY, 1981, 1986), e mais precisamente o Parâmetro do Sujeito Nulo. Nossa pesquisa propõe-se a testar uma hipótese não prevista em outros estudos realizados até então, a de que os alunos que cursam seus períodos finais na UNILAB podem apresentar, em decorrência do contato intenso com o português brasileiro, alguns aspectos que se diferenciam do uso da norma europeia do português.

É notório e amplamente defendido o conceito que caracteriza o português europeu – língua do colonizador dos países que abordamos aqui, dentre outros – como uma língua de sujeito nulo consistente, em decorrência da riqueza de sua morfologia de flexão verbal, o que, ao contrário do que se constata na atual sincronia do português brasileiro, lhe confere uma alta frequência de sujeitos nulos (DUARTE, 1995, 2000; GALVES, 2001). Logo, analisar a natureza da morfologia de flexão verbal no português de Guiné Bissau e no português de São Tomé e Príncipe é um dos objetivos deste trabalho e para tal análise foram selecionados alunos que estão nos primeiros e últimos períodos de diferentes cursos de graduação da UNILAB. Por conseguinte, dados de fala e escrita desses alunos foram coletados e analisados tendo como hipóteses: a) nos primeiros anos, pode ser verificada uma maior assimetria entre guineenses e são-tomenses: os primeiros tendem a usar menos sujeitos nulos na escrita e na fala que os segundos em virtude de utilizarem mais o crioulo guineense, uma língua que não permite sujeitos nulos e b) nos últimos anos, embora seja esperado um uso mais próximo ao português europeu do português guineense e do português são-tomense, conforme previsto em outras pesquisas, espera-se, a partir do contato intenso com o português brasileiro, encontrar restrições nessas duas variedades africanas do português para a distribuição de sujeitos nulos e plenos na língua falada e escrita não verificadas em português europeu.

A partir de leituras realizadas de trabalhos sobre o Parâmetro do Sujeito Nulo, que apontavam para uma mudança na gramática do português brasileiro no que diz respeito à verificação de uma alta frequência de sujeitos plenos, que, ao contrário do que acontece com o português europeu, se apresenta como consequência do enfraquecimento da morfologia de flexão verbal motivada pela inserção do pronome você e da expressão pronominal a gente (DUARTE, 1993, 1995), surgiu a motivação para fazer esse trabalho. Nossa motivação também se relaciona com a oportunidade de ampliar o estudo de trabalhos anteriores abarcando outras variedades não-europeias do português, bem como com a escassez, verificada até o presente momento, de pesquisas que abordem o Parâmetro do Sujeito Nulo nas variedades do português falado em São Tomé e Príncipe e Guiné Bissau.

Em geral, os estudos concentram-se em apenas uma modalidade de uso da língua, contudo, nossa pesquisa se dedicou ao trabalho comparativo entre a língua falada e a língua escrita de um mesmo estudante (guineense e são-tomense) da UNILAB. Desta forma, esta pesquisa se apresenta como relevante e inovadora, visto que não se resume à análise de apenas uma modalidade de uso da língua. Essa comparação fez-se necessária, devido a possibilidade da ocorrência de assimetria entre a língua falada e a língua escrita. Esta

ocorrência fomentaria a hipótese de que restrições encontradas em português brasileiro para a distribuição de sujeitos nulos seriam mais frequentes na língua falada e escrita dos estudantes dos últimos períodos da UNILAB. No que tange aos alunos dos primeiros períodos, os estudantes guineenses apresentam maior tendência de reprodução de sujeitos plenos em decorrência da predominância do uso do crioulo guineense, ao contrário do que se observa em São Tomé, onde a maioria da população já tem o português como primeira língua (HAGEMEIJER, 2009).

De modo geral, com este trabalho esperamos contribuir para desenvolvimento de outros estudos no campo da sintaxe comparativa no âmbito do parâmetro do sujeito nulo, sobretudo, através do contraste entre as modalidades da língua falada e escrita de variedades não-europeias do português, uma perspectiva ainda pouco difundida entre os estudiosos da área.

Para realizar nossas análises usaremos a abordagem hipotético-dedutiva e, como métodos de procedimento, o estatístico e o comparativo: o primeiro, para a realização da análise quantitativa e o segundo, para estabelecer a comparação entre as variedades não-europeias do português selecionadas para a análise no que se refere ao sujeito nulo.

Para melhor compreensão do nosso trabalho, organizamo-lo da seguinte maneira: no capítulo introdutório, discorreremos sobre o nosso objeto de estudo que se relaciona ao PSN, bem como sobre estudos realizados sobre esse parâmetro em variedades do português, a fim de contextualizarmos esse objeto. Além disso, apresentamos os objetivos e as hipóteses da pesquisa. No segundo capítulo, abordaremos o quadro teórico que norteia a análise dos dados. Centraremos nossa atenção nos pressupostos da gramática gerativa e no Modelo de Princípios e Parâmetros, em especial, no que se refere ao PSN e, por fim, discorreremos sobre a aquisição da língua – L1 e L2 – à luz desse modelo teórico. No terceiro capítulo, detalharemos os procedimentos metodológicos utilizados para realizarmos esta pesquisa, levando em conta, por exemplo, o método empregado, a natureza da pesquisa, os corpora elaborados, as variáveis selecionadas e o tratamento quantitativo a que os dados se submeteram. O quarto capítulo, por sua vez, apresentará os resultados de nossa pesquisa, tomando por base a comparação entre o PGB e o PST na língua falada e escrita dos estudantes que estão nos primeiros e últimos anos da UNILAB. Para tanto, a análise linguístico-comparativa será feita com base no Modelo de Princípios e Parâmetros. Por fim, abordaremos nas considerações finais os principais resultados obtidos, levando em conta as semelhanças

e/ou diferenças encontradas entre o PGB e o PST no que se refere ao PSN. E ainda, são levantados alguns questionamentos para futuras investigações.

Finalmente, de acordo com os resultados obtidos, constatamos que, em relação à distribuição do sujeito nulo, há diferenças entre as variedades africanas analisadas, conforme tínhamos previsto inicialmente. O português guineense apresentou índices menores em relação ao uso do sujeito nulo quando contrastado com o português são-tomense, enquanto que esse apresentou resultados que vão ao encontro do português europeu, no que concerne ao uso do sujeito nulo. Percebemos também que a interferência do português brasileiro se deu de maneira mais expressiva na modalidade falada da língua, principalmente no português guineense, assim, identificando a língua escrita como um contexto de monitoramento para o não preenchimento da posição de sujeito.

2 O PANORAMA DO ESTUDO

2.1 DELIMITAÇÃO DO OBJETO DE ESTUDO

Alguns estudos recentes têm demonstrado semelhanças morfo-sintáticas entre a variedade do português brasileiro e as variedades africanas do português, a exemplo, o trabalho de Petter (2009, p. 218) que trata de aspectos históricos e linguísticos (contato linguístico, semelhanças no léxico, morfologia das línguas) comuns ao português angolano (PA), moçambicano (PM) e brasileiro (PB). A autora afirma que “[a] semelhanças dos fenômenos analisados em sincronia (desvios, erros, particularidades, no PA e PM, e variação, no PB) constituem dados bastante eloquentes em favor da hipótese de um *continuum* afro-brasileiro de português”, distanciando-se assim da língua do colonizador, o português europeu (PE), e constituindo-se em variedades não-europeias do português. Assim, lidamos com variedades – africanas e brasileira – desse português europeu.

Tanto no Brasil quanto na África, verifica-se um contexto de intenso multilinguismo, conforme defende Mattos e Silva (1998, p. 16): “[c]ontrariamente ao que ocorre no português europeu, em todas as outras áreas lusófonas, o português se encontra em situações complexas e diversas de multilinguismo e aí está incluído o Brasil.”. Na África, por exemplo, o português, apesar de ser a língua oficial em cinco países (Cabo Verde, Guiné Bissau, São Tomé e Príncipe, Moçambique e Angola), convive com outras línguas. Para entendermos um pouco esse complexo multilinguismo, segue-se um quadro em que são apresentadas as línguas mais faladas que convivem com o português em cada país:

Quadro 1 - Língua(s) mais usada(s) e língua oficial em cinco países africanos. Extraído de Oliveira e Silva 2018

PAÍS	LÍNGUA MAIS FALADA	LÍNGUA OFICIAL
Angola¹	Línguas do grupo bantu (família benue-congo ocidental): a) <i>umbundo</i> (língua materna falada por cerca de um terço dos angolanos); b) <i>kimbundo</i> (língua	<i>Português</i> : ✓ primeira ou segunda língua: falada por 71% dos angolanos ³

¹ Conforme Petter 2015, em Angola e Moçambique, falam-se línguas bantas, ao passo que, nos demais países lusófonos da África, os crioulos. Essa classificação, dá-se pela relação com a língua de origem e pelo posicionamento territorial geográfico, os crioulos surgem em regiões de ilhas e está muito atrelado ao processo de colonização. É importante ressaltar que os crioulos de Cabo-Verde, São Tomé e Príncipe e Guiné Bissau são

	falada pela quarta parte da população); c) <i>kikongo</i> ; d) <i>côkwe</i> , e) <i>ganguela</i> e f) <i>cuanhama</i> . ²	
Moçambique	Línguas do grupo bantu (família benue-congo): a) <i>macua (emakhuwa)</i> (língua materna falada por 26,3%); b) <i>changana</i> (língua materna falada por 11,4%) e c) <i>elomwe</i> (língua materna falada por 7,9%) ⁴ .	<i>Português</i> : ✓ cerca de 50% falam português (10,7% dos falantes consideram-no língua materna e 39,7%, como segunda língua. (GONÇALVES, 2015)
Cabo Verde	Língua materna: <i>crioulo caboverdiano</i> (quase todos os caboverdianos falam-na (926.078 falantes) ⁵	<i>Português</i> ✓ 87% falam português ⁶
São Tomé e Príncipe	Línguas crioulas: a) <i>santomé/ forro</i> (língua de São Tomé): 72,4% (maiores de cinco anos) ⁷ ; b) <i>angolar</i> (língua dos Angolares): 5,3% da população ⁸ e c) <i>lung'ie</i> (língua do Príncipe): 2,4 % (mais de cinco anos de idade) ⁹ .	<i>Português</i> ✓ 98,4% falam português ¹⁰
Guiné-Bissau	crioulo guineense (língua materna falada por cerca de 80% da população) (OLIVEIRA; BAILO; INJAI, 2013)	<i>Português</i> ✓ Cerca de 13% da população usam-na como língua segunda ou terceira. (COUTO; EMBALÓ, 2010)

Observando o quadro 1, percebemos que, entre São Tomé e Príncipe e Guiné Bissau, há uma assimetria: enquanto naquela o português é frequentemente falado, nesta predomina o

crioulos de base lexical portuguesa com características fonético-fonológicas e morfossintáticas de línguas nativas.

³ https://pt.wikipedia.org/wiki/Angola#cite_note-57. Informação contida no *Observatório da Língua Portuguesa*. Consultado em 30 de março de 2016.

² <https://pt.wikipedia.org/wiki/Angola#L.C3.ADnguas>

⁴ Instituto Nacional de Estatística de Moçambique Línguas mais faladas em Moçambique, de acordo com o censo de 1997

⁵ https://pt.wikipedia.org/wiki/Crioulo_cabo-verdiano

⁶ Informação extraída do Observatório da Língua Portuguesa: <https://observalinguaportuguesa.org/category/dados-estatisticos/>

⁷ Censo de 2001. Informação contida em Hagemeijer (2009).

⁸ Conferir Lorenzino (1998)

⁹ Censo de 2001. Informação contida em Hagemeijer (2009).

¹⁰ Informação extraída do Observatório da Língua Portuguesa: <https://observalinguaportuguesa.org/category/dados-estatisticos/>

crioulo guineense. Segundo Hagemijer (2009, p. 19-20), São Tomé e Príncipe é o “único país da África de língua portuguesa onde a maioria da população tem actualmente o Português como primeira língua, havendo assim condição para a emergência de uma nova variedade”.

Estudos ainda revelam que quanto maior é o uso de outra língua (como é o caso do crioulo guineense), maior é o afastamento da variedade africana do português em relação à variedade europeia do português. Segundo Duarte (2006), é apenas, no século XIX, que se instala o ensino da língua portuguesa em Guiné Bissau, sendo o crioulo guineense utilizado nos mais diversos domínios, encontrando-se estruturalmente bem mais estabilizado do que a língua do colonizador.

Ademais, falantes com nível superior apresentam um uso linguístico mais próximo do PE. Brandão e Vieira (2012), ao analisarem o domínio da concordância verbal e nominal na fala de são-tomenses, observam que a regra linguística nesse domínio é variável e que falantes com nível superior e que usam menos o Santomé, crioulo de São Tomé, tendem a produzir mais concordância.

Em se tratando de outro aspecto sintático relacionado ao Parâmetro do Sujeito Nulo (PSN) conforme Chomsky (1981, 1986 e seguintes), em que há línguas em que a posição sujeito é obrigatoriamente preenchida por sujeitos plenos e outras em que essa posição pode ser preenchida por nulos, há muitos estudos comparativos entre o PB e o PE, Duarte (1995); Kato e Negrão (2000); Galves (2001); Magalhães (2004); Silva (2004); Kato e Duarte (2014a, 2014b). Esses estudos, mostram evidências de que, no PB, sujeitos plenos são mais frequentes, ao contrário do PE por razões a serem discutidas na seção 1.3.1. No entanto, até onde temos verificado, estudos sobre esse parâmetro ainda são escassos nas variedades africanas do português.

Como é sabido, Chomsky (1981, 1986), ao defender, a partir do Modelo de Princípios e Parâmetros, que a Gramática Universal (GU) consiste de um conjunto de propriedades gramaticais compartilhadas por todas as línguas naturais (os Princípios) e dos parâmetros definidos como opções dentro da própria GU que devem dar conta das diferenças entre as línguas, propõe o PSN (também chamado de Parâmetro Pro-drop), através do qual as línguas se subdividem em: línguas de sujeito nulo (*línguas pro-drop*) por permitirem a não-realização fonética do sujeito e línguas de sujeito não-nulo (*línguas não-pro-drop*) que obrigam essa realização.

Assumindo que todas as línguas naturais projetam a posição de sujeito em contextos finitos (o chamado “Princípio da Projeção Estendida”), Chomsky (1981, 1982) e Rizzi (1982) verificaram que, em línguas pro-drop, por ser a morfologia rica, sujeitos nulos são

identificados pela morfologia de flexão verbal através dos morfemas número-pessoais, ao contrário de línguas com morfologia pobre. Sendo assim, em línguas que possuem uma morfologia flexional rica, o PSN é marcado positivamente, a exemplo do espanhol (cf. (1a) e do italiano (cf. 1b)), ao passo que, em línguas com morfologia flexional pobre, é marcado negativamente, como é o caso do inglês (cf. (2a)) e (2b)):

- (1) a. *pro* Hablo español
 b. *pro* Parlo spagnolo

- (2) a. *I* speak Spanish
 b. *Je* parle espagnol

Como é observado nas sentenças acima, (1a) e (1b) representam línguas *pro-drop*, cuja posição sujeito é ocupada pela categoria vazia *pro* (nomeadamente, o sujeito nulo), valendo referirmos que, nessas línguas, é possível que sujeitos plenos ocorram quando, por exemplo, recebem acento enfático, contrastivo. Em se tratando de (2a) e (2b), o sujeito deve ser realizado foneticamente (nomeadamente, o sujeito pleno), pois a morfologia de flexão verbal é incapaz de identificá-lo.

Discutindo o PSN em línguas como o chinês, Huang (1984, 1989) observou que, embora não possuam morfologia de flexão rica, sujeitos nulos são produzidos quando mantêm referência conjunta com um referente já dado no domínio do discurso, um fato que fez a proposta de Chomsky (1981) e Rizzi (1986) ser ampliada da seguinte maneira: em línguas como o italiano e espanhol, o identificador do sujeito nulo é realizado por um “controlador interno”, a saber, a morfologia de flexão verbal, ao passo que, em línguas como o chinês e o japonês, o sujeito nulo é identificado por um antecedente que se conecta com a categoria vazia por meio de uma relação de correferência. Diz-se, então que a identificação do sujeito nulo está ligada a um tópico no discurso, conforme verificamos em (3):

- (3) Zhangsan shuo [*e* hen xihuan Lizi].
 (“Zhangsan disse que (ele) gostou de Lizi”) (Huang, 1989, p. 187)

De acordo com Huang (1989), é possível fazer duas observações para a oração em (3):
 a) o sujeito nulo da oração encaixada (representado por *e*) pode estar ligado ao sujeito da

matriz e b) esse sujeito pode referir-se a outra pessoa, cuja referência é compreendida por meio do discurso.

No decorrer de algumas décadas, várias propostas de análise foram feitas para o PSN e destacaremos brevemente aqui a proposta feita por Holmberg, Nayudu e Sheehan (2009) e Roberts e Holmberg (2010) para a formulação de quatro sistemas de línguas pro-drop. Vejamos:

Quadro 2 - Sistema de sujeitos nulos com base nos estudos de Holmberg, Nayudu e Sheehan (2009) e Roberts e Holmberg (2010)

Línguas de sujeito nulo consistente	Línguas que possuem morfologia de flexão verbal rica e que, portanto, licenciam sujeitos nulos. Exemplos: o espanhol, o italiano e o português europeu
Línguas de sujeito nulo radical	Línguas que não possuem morfologia de flexão verbal, sendo o sujeito nulo produzido sob restrições discursivas. Exemplo: o chinês e o japonês.
Línguas de sujeito nulo parcial	Línguas em que o sujeito nulo é submetido a contextos estruturais específicos. Exemplos: o português brasileiro e o hebraico.
Línguas de sujeito nulo expletivo	Línguas em que apenas o sujeito expletivo é nulo, não havendo possibilidade de sujeitos nulos referenciais. Exemplo: o alemão.

Com a criação desse sistema, houve a possibilidade de agrupar as línguas em conjuntos mais específicos, de acordo com as características que eram apresentadas em relação à marcação positiva para o parâmetro do sujeito nulo. Assim, não se tem apenas a binaridade da marcação: positiva ou negativa para o parâmetro, relacionando-o à morfologia flexional rica. Ele pode ser marcado positivamente, mesmo que não apresente essa morfologia por encaixar-se dentro das propriedades dos outros sistemas de língua pro-drop.

Como dissemos anteriormente, muitos são os estudos realizados sobre o PSN durante esses anos. Fizemos aqui um breve resumo de algumas pesquisas para basearmos a nossa

próxima discussão que trata do parâmetro pro-drop em diferentes variedades do português. No capítulo dois, faremos uma abordagem mais aprofundada sobre o PSN.

Como é sobejamente defendido, o PE, língua do colonizador nos cinco países africanos (cf. quadro 1 acima) cuja independência ainda é recente, é uma língua de sujeito nulo consistente, pois há uma altíssima frequência de sujeitos nulos em decorrência da riqueza de sua morfologia de flexão verbal, ao contrário do que se observa na atual sincronia do PB, de acordo com Duarte (1995, 2000); Galves (2001). Hoje podemos dizer que o PB é uma língua de sujeito nulo parcial, de acordo com a proposta de Holmberg, Nayudu e Sheehan (2009) e Roberts e Holmberg (2010), conforme será apresentado na seção 1.3.

Diante do exposto e voltando nossa atenção às variedades africanas do português, no entanto, verificamos que ainda são escassos estudos sobre o PSN. Sobre o PA, Oliveira e Santos (2007) verificam que se aproxima do PE, pois sujeitos nulos são mais frequentes, o que vai também na direção dos resultados obtidos por Bravin dos Santos (2010) para o PM. No que concerne ao português cabo-verdiano (PCV), Alexandre, Duarte e Santos (2013) assumem que é uma língua de sujeito nulo parcial, aproximando-se assim mais do PB que do PE. Com relação ao português são-tomense (PST), não encontramos, até o presente momento, uma pesquisa que trate especificamente e de forma detalhada da relação entre o PSN e esta variedade do português. Hagemeyer (2007), em sua tese, aborda o PSN no crioulo são-tomense, afirmando que essa é uma língua de sujeitos não nulos. Conforme mencionado anteriormente, Brandão e Vieira (2012) verificam que o PST é uma língua cuja regra de concordância verbal apresenta-se como variável, fato que possibilita uma aproximação com o PE. Além disso, por ser a língua mais falada pela grande maioria da população (98,4%) e sendo o santomé uma língua de sujeito não nulo – mas que é falada por uma minoria da população –, supomos que o PST tenha comportamento muito próximo ao PE, pois sujeitos nulos podem ser muito produtivos.

No tocante ao português de Guiné Bissau (PGB), pesquisas também têm sido escassas. Conforme o quadro 1, apenas 13% da população guineense falam o PGB e 80% têm o crioulo guineense como língua materna cujo parâmetro do sujeito nulo tem valor negativo, conforme Castro (2013). Assumindo esses fatos, supomos aqui que o PGB, ao contrário do PST, pode apresentar mais sujeitos plenos pela maior influência da língua materna, que é a língua crioula. Lucchesi e Baxter (2009) trabalham com o conceito de Transmissão Linguística Irregular para designar os processos históricos de contato entre povos falantes de línguas tipologicamente diferenciadas, em decorrência do colonialismo europeu na África e em outros continentes. De acordo com os autores, o processo de transmissão linguística irregular pode

levar à formação de uma língua historicamente nova, denominada língua crioula, como aconteceu tanto em São Tomé quanto em Guiné Bissau, ou à formação de uma nova variedade histórica da língua de superstrato, como vimos, os estudos de Hagemeyer (2007, 2009) já apontam o PST como uma variedade da língua portuguesa. Essa nova variedade histórica da língua de superstrato não deixa de apresentar os processos de variação e mudança movidos pelo contato entre as línguas.

A partir da teorização feita acima, analisaremos os dados de língua escrita e falada dos informantes dos primeiros períodos em contraste com os últimos períodos, pois acreditamos que ao analisar os dados dos informantes recém-chegados ao Brasil não encontraremos influência do PB, em contrapartida, consideramos encontrar influências do PB nos dados dos informantes dos últimos períodos, devido ao tempo do contato linguístico com o português brasileiro. No que concerne às variedades não-europeias estudadas, esperamos encontrar maior proximidade do PST em relação ao PE, pois, como vimos, o português é falado pela maior parte da população são-tomense, enquanto que no PGB supomos encontrar maior proximidade com o crioulo guineense, pois esta é considerada língua materna e falada por grande parte da população guineense. Por fim, acreditamos que os dados de um mesmo informante podem ter assimetrias entre a fala e a escrita, devido a um maior policiamento linguístico dos informantes no texto escrito, por isso, trabalharemos com as duas modalidades da língua, assim, possibilitando uma análise mais completa.

Diante do exposto, centraremos, portanto, nossa atenção na comparação entre falantes guineenses e são-tomenses que moram no Brasil e estudam na Universidade de Integração Internacional da Lusofonia Afro-Brasileira (UNILAB), localizada na cidade de Redenção-CE, tomando por base a distribuição de sujeitos nulos e plenos em dados orais e escritos por eles produzidos, tendo em mente possíveis interferências do PB. Para tanto, a análise estará embasada no Modelo de Princípios e Parâmetros, Chomsky (1981, 1986), em específico, no Parâmetro do Sujeito Nulo.

Vale referirmos que o trabalho comparativo a ser realizado entre o PGB e o PST é relativamente pioneiro e trará importantes resultados no que se refere a semelhanças e/ou diferenças entre essas variedades não-europeias do português em relação à distribuição dos sujeitos nulos e plenos e às razões que estão na base das possíveis diferenças, não esquecendo da importante história de contato linguístico dessas línguas com línguas crioulas, conforme será abordado no próximo tópico.

Além disso, um aspecto inovador e relevante deste estudo é a comparação que se pretende realizar entre a língua falada e a língua escrita de um mesmo estudante (guineense e

são-tomense) da UNILAB, pois, em geral, os estudos concentram-se em apenas uma modalidade de uso da língua, a saber: a língua falada. Essa comparação faz-se necessária, tomando por base que possa ocorrer uma possível assimetria entre a língua falada e a língua escrita.

Em linhas gerais, esperamos contribuir com os estudos no campo da sintaxe comparativa no âmbito do PSN, sobretudo, através do contraste entre a língua falada e escrita de variedades não-europeias do português, uma perspectiva ainda pouco difundida entre os estudiosos da área.

2.2 O CONTATO LINGUÍSTICO E AS VARIEDADES AFRICANAS DO PORTUGUÊS

Sabemos que, historicamente, uma língua de contato é o resultado de uma situação linguística caracterizada pela presença no mesmo espaço de muitas línguas distintas, conforme Thomason (2001). Essa realidade foi e ainda hoje é representada no continente africano. Os habitantes desse continente não viviam isolados. Mesmo antes da chegada dos europeus, eles interagiam entre si por várias razões: sociais, econômicas, políticas e culturais, ou ainda, quando precisavam migrar para novas regiões. Ao contrário do que podemos pensar, antes mesmo da colonização, já existia uma enorme variedade linguística. Havia mais de duas mil línguas e esse fato não era um impedimento na comunicação, pois os habitantes estavam acostumados com as línguas de seus vizinhos e com as relações comerciais. Segundo Petter (2015, p. 193), é a partir da colonização que as línguas africanas existentes começam a mudar, com a interação entre a língua portuguesa e as línguas africanas passa a existir uma nova organização linguística:

O advento do colonialismo, no entanto, ao introduzir novos atores e novas línguas – não africanas –, vai provocar, entre outros efeitos, um rearranjo do relacionamento entre as línguas em presença. Instaura-se, então, uma nova dinâmica de convivência linguística com a superposição de um elemento crucial: a atitude das nações colonizadoras em relação às línguas locais.

Segundo a autora, a partir de 1884 quando a África começa a ser dividida, cada nação colonizadora procurou adotar sua língua materna como língua oficial do país colonizado, com a finalidade de facilitar a administração e a educação assim como ocorreu no Brasil. Houve

nação colonizadora que integrou sua língua ao sistema educacional – possibilitando o ensino da língua africana e da língua do colonizador (a Grã-Bretanha introduziu as línguas africanas nos anos iniciais da escolarização, para posteriormente ensinar a língua inglesa) -, houve outras, como Portugal, que deram pouca importância ao processo de introdução das duas línguas ao nativo, fazendo com que o ensino fosse exclusivamente em português. Assim, era possível distinguir o africano ‘civilizado’, aquele que havia aprendido a língua do colonizador. É importante salientarmos que, na época em que o continente africano estava sendo dividido, o Brasil já era independente de Portugal, o que pode ser uma das justificativas do avanço das políticas linguísticas brasileiras, ao contrário do que ocorre nos países africanos colonizados por Portugal cujas políticas ainda estão ganhando espaço e implementação, como as pesquisas que são feitas para a implementação de um ensino bilíngue em Guiné Bissau, por exemplo.

No continente africano, existe uma espécie de “competição” linguística na qual o falante faz, de forma consciente ou intuitiva, escolhas entre as línguas que estão em contato em seu meio comunicativo. Essas escolhas são feitas de acordo com as funções que as línguas e suas variedades desempenham em cada contexto comunicativo. Por essa razão, faz-se necessária a promoção de políticas linguísticas que garantam a coexistência dessas línguas de forma organizada. Sentindo a necessidade de orientar a convivência entre as línguas, alguns países africanos já desenvolvem planejamentos linguísticos com essa finalidade.

O modo pelo qual se dá o uso da língua em cada um dos contextos comunicativos é decorrente de uma demanda social, contudo, esse uso também pode decorrer das diretrizes estabelecidas por um planejamento no qual se relacionam questões políticas e sociais com os estudos linguísticos. Dentre os aspectos que baseiam e definem essas diretrizes estão os contextos de uso da língua, a sua categorização como língua nacional ou oficial e as ações que podem ser aplicadas para que determinada língua se difunda no sistema educacional. O planejamento linguístico se faz necessário como uma ferramenta capaz de solucionar possíveis problemas de comunicação oriundos da relação entre as línguas, pois existe um movimento que aponta para a flexibilização do uso e da função das línguas na sociedade, possibilitando o uso dessas línguas em diferentes contextos – porém, essa mudança demanda tempo e não seria possível sem o apoio e o incentivo do governo. Logo, uma das tarefas primordiais da linguística nesse contexto é proporcionar a padronização ou gramatização das línguas locais com a elaboração de dicionários e gramáticas. O processo de gramatização do Brasil aconteceu logo após a independência do país, enquanto na África esse processo ainda

se revela muito lento, segundo Sani (1996, p. 23), o interesse nos estudos sobre bilinguismo progrediu de 1950 a 1980, com alguns trabalhos que buscavam, dentre outros objetivos, “explicar a convivência criada pelo trabalho em um meio linguístico majoritário produzindo uma variedade de línguas em uma comunidade bilíngue”. Contudo, embora exista o interesse por parte dos pesquisadores em melhorias para um contexto de ensino bilíngue, há uma trava das autoridades governamentais que insistem em um ensino exclusivamente feito em língua portuguesa.

Depois do exposto, apresentaremos, nas subseções 2.2.1 e 2.2.2), um breve panorama da situação sociolinguística de São Tomé e Guiné Bissau com a finalidade de compreendermos o contato linguístico do português com as línguas maternas desses países.

2.2.1 A situação sociolinguística de São Tomé e Príncipe

O povoamento de São Tomé e Príncipe aconteceu por volta de 1493, mas é a partir de 1520 que escravos de vários países da África foram trazidos para trabalhar na cultura da cana-de-açúcar. Nesse contexto, é fato que, em termos numéricos, os africanos superavam os colonizadores. Por esses motivos, a fase de colonização permitiu um contato intenso entre portugueses e escravos mediante à urgência da comunicação entre eles. Segundo Hagemeyer (2009), do contato desses escravos com os senhores portugueses, surgiu um *pidgin*¹¹, que, por sua vez, deu lugar a línguas crioulas cujo “léxico [é] formado principalmente pela língua do ex-colonizador e [possui] uma gramática constituída pela reestruturação de traços das várias línguas faladas nas regiões onde surgiram.” (PETTER, 2015, p. 211). Nesse sentido, para a formação dos crioulos nas ilhas de São Tomé e Príncipe, o português atua como língua de superstrato, ou seja, aquela que “dá” o léxico e, portanto, são considerados crioulos de base lexical portuguesa (ex: Santome ou Forro, Lung’ie ou Principense, Angolar ou Ngola) e as línguas nativas atuam como línguas de substrato, ou seja, “línguas que estiveram em contacto com o Português [por exemplo] na fase de formação dos crioulos” (PEREIRA, 2007, p. 49). No caso do Santome, a língua de substrato é o Edo, língua do tronco nígero-congolês da família benue-congo (HAGEMEIJER, 2007, 2009).

¹¹ Os *pidgins* possuem “uma estrutura restrita e desempenham uma função especializada; normalmente serviram para as trocas comerciais, e foram usados por não nativos como *língua franca* por falantes que conservaram suas línguas vernáculas. A maior parte de seu vocabulário vem da língua dominante, normalmente europeia, e a estrutura gramatical vem de uma ou de várias línguas vernáculas (africanas, no caso).” (PETTER, 2015, p. 211)

Porém, mesmo com o aparecimento e permanência das línguas crioulas nas ilhas de São Tomé e Príncipe, o português ainda é a língua mais falada. A partir de sua independência de Portugal em 1975, houve uma maior mobilidade social e acesso ao ensino e aos meios de comunicação na língua oficial - o português – e a ausência de políticas orientadas para a manutenção e valorização das línguas crioulas, tornando-as minoritárias nas ilhas e reservadas a contextos informais e à modalidade oral.

De acordo com Hagemer (2009, p. 19-20), São Tomé e Príncipe: é “o único país da África de língua portuguesa onde a maioria da população tem actualmente o Português como primeira língua, havendo assim condição para a emergência de uma nova variedade”. E ainda,

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

Tomando por base que há dados do PST que se aproximam e se afastam da norma europeia do português, a exemplo, temos alguns trabalhos: Silveira (2013) investiga a realização dos ditongos orais no português de São Tomé e Príncipe e constata que essa variedade difere do sistema linguístico da variedade considerada padrão, o português europeu; Gonçalves (2016) analisa as construções ditransitivas no português de São Tomé em comparação com o PE, considerando o contato linguístico com outras línguas africanas; entre outros. Investigaremos neste estudo se, no campo morfossintático relacionado ao PSN, encontraremos mais proximidade ou não com o uso do português europeu, considerando por exemplo, as modalidades de uso da língua: língua falada e língua escrita e a comparação feita entre os dados dos informantes chegados recentemente ao Brasil (alunos cursando os primeiros períodos) em confronto com os dados dos informantes que estão no Brasil a mais de três anos (alunos cursando os últimos períodos).

Ao analisarem a concordância nominal e verbal no PST, Brandão e Vieira (2012, p. 1055) observam:

[e]mbora seja, de fato, a escolaridade o fator que faz mudar as tendências relativas ao fenômeno (variável da concordância verbal) [...] está a maior ou menor utilização da língua crioula. Só declararam que usam frequentemente

o crioulo informantes que cursaram até o ensino fundamental, nível de escolaridade em que se registra maior cancelamento da marca de número. De fato, o prestígio que recebe o Português em terras são-tomenses e o maior ou menor propósito de dominar essa língua, o que pode se tornar possível graças ao acesso à escolaridade, trazem por consequência o maior ou menor domínio das regras canônicas de concordância.

De acordo com as autoras, a variável nível de escolaridade associa-se ao maior ou menor uso da língua crioula: quanto menos escolarizado é o falante, mais o crioulo é usado e mais ausência de concordância verbal e nominal é verificada. Esse aspecto não será abordado em nosso trabalho, pois, não temos o nível de escolaridade como uma variável (todos os informantes estão no mesmo nível de escolaridade), em contrapartida, analisaremos o tempo que os estudantes estão no Brasil e se há uma influência do PB na língua escrita e falada desses estudantes. Embora não tenhamos a variável nível de escolaridade em nosso trabalho, reconhecemos sua importância nos resultados de outras pesquisas, a saber: Brandão e Vieira (2012) que ao analisarem a concordância verbal no PST e PB verificam que a escolarização é o fator mais relevante em ambas as variedades, interferindo sensivelmente nos resultados. Bazenga (2016) afirma que o nível de escolarização superior é um fator determinante na concordância de P6, atingindo 80% no nível 1 e 90% no nível 3 de escolaridade, entre outros estudos.

2.2.2 A situação sociolinguística de Guiné Bissau

Guiné Bissau foi a primeira colônia portuguesa a ter sua independência reconhecida por Portugal. Porém, essa independência foi conquistada de maneira dolorosa. Seu primeiro líder, Amílcar Cabral, tentou uma negociação pacífica, que foi rejeitada por Portugal. Em 1959 começa a luta armada e apenas em 1974 é que Guiné Bissau tem sua independência reconhecida.

Contudo, conquistar a independência não foi sinônimo de êxito. Já havia no país a falta de infraestrutura e os antigos colonizadores deixaram os cofres públicos vazios, impedindo os investimentos necessários em áreas como: saúde, segurança e educação. Naquela época, o índice de analfabetismo era de quase 99% e infelizmente não houve uma mudança até os dias atuais. Embora seja um território pequeno, há mais de vinte línguas sendo faladas em Guiné Bissau. Desde o tempo da colonização, foi instaurado o português como língua oficial, entretanto, apenas 13% da população usa-a como segunda ou terceira língua. A

língua nacional é o crioulo guineense – língua materna – que é falada por cerca de 80% da população. As demais línguas são nativas do continente africano e, embora existam em grande número, não passaram por um processo de gramatização que depende, como é sabido, de política e planejamento linguístico, que devem ser desenvolvidos por interesse governamental. Segundo Couto e Embaló (2010, p. 42):

Como a alfabetização e todo o ensino posterior se dá numa língua estrangeira, os resultados são calamitosos. Assim, somente um em cada 500 alunos transita da primeira à décima primeira classe sem nenhuma repetição. No nível elementar, apenas um em 400 chega ao sexto ano com sucesso. 41% dos alunos inscritos na primeira classe não são admitidos na segunda. Isso tudo levando em consideração que apenas 40% das crianças guineenses se matriculam em alguma escola. Portanto, não é de se admirar que a taxa de analfabetismo é de 86%.

Esses números exorbitantes acontecem, pois, desde a época da colonização, o ensino é feito em português, língua falada por uma minoria da população. Por isso, embora seja oficialmente proibido nas escolas, conforme visto em Fonseca (2012), há em Guiné Bissau o Decreto-lei – nº 7/2007 de 12 de Novembro de 2007 –, que obriga à utilização da língua portuguesa em todas as instituições públicas, nomeadamente nas escolas e especificamente dentro da sala de aula e no recinto escolar, mesmo assim, alguns professores precisam iniciar o ensino – nos primeiros anos – com o crioulo guineense, pois, ao usarem o português não são entendidos pelas crianças, conforme Couto e Embaló (2010).

Ainda, segundo Couto e Embaló (2010), a maioria dos guineenses está em comunidades em que a língua materna é uma língua africana minoritária e, como é comum, sua aquisição dá-se no meio familiar, por meio do contato com seus pais e as pessoas que os cercam. Para grande parte dos guineenses, não há relação de pertencimento para com a língua portuguesa, visto que a língua afetiva é o crioulo guineense, o que claramente dificulta o aprendizado do português. Sabendo disso, reconhecemos o quão difícil é a convivência do português de Guiné Bissau com as línguas africanas.

2.3 SOBRE O PARÂMETRO DO SUJEITO NULO EM VARIEDADES DO PORTUGUÊS

2.3.1 O parâmetro do sujeito nulo no português europeu

Tem sido sobejamente assumido que o PE é uma língua pro-drop prototípica, conforme Barbosa (1996), Barbosa, Duarte e Kato (2001), Duarte (2000), ou seja, uma língua de sujeito nulo consistente, à semelhança do espanhol e do italiano, tomando por base o sistema de línguas de sujeitos nulos proposto por Holmberg, Nayudu e Sheehan (2009). Sabemos que nessas línguas o licenciamento de sujeitos nulos está associado à morfologia de flexão verbal rica que, por meio dos traços-*phi* (número e pessoa), permitem recuperar os traços *phi* do sujeito nulo.

Roberts (1993) propõe o seguinte: um paradigma de flexão verbal que pode possuir uma desinência zero e um sincretismo seria “funcionalmente rico”, sendo capaz de licenciar e identificar os sujeitos nulos. Observemos o comportamento do PE:

Quadro 3 - Paradigma de flexão verbal extraído de Galves (2001, p. 103)

Português Europeu	
1ª pessoa do singular	(eu) canto
2ª pessoa do singular	(tu) cantas
3ª pessoa do singular	Você/Ele canta
1ª pessoa do plural	(Nós) cantamos
2ª e 3ª pessoas do plural	(Vocês/eles) cantam

Como pode ser observado no quadro acima, mesmo no PE não se faz mais comum o uso do pronome *vós*¹² para a segunda pessoa do plural. É essa situação que faz a língua apresentar dois sincretismos: a desinência zero para a 3ª pessoa do singular e a desinência *-m* para a 2ª e 3ª pessoa do plural. Galves (2001) acredita que esse sincretismo – como é ocasionado pelo acréscimo de formas de tratamento (*você-vocês*) – não deve comprometer a existência funcional do paradigma.

Para corroborar sua análise, Duarte (1995), que estudou como a perda do princípio “Evite Pronome” vem acontecendo no português brasileiro, apresenta um quadro, sobre o

¹² O uso do pronome *vós* é dispensável ao tratarmos do dialeto lisboeta. No Porto, ainda se observa esse uso.

português europeu, que mostra que todas as pessoas gramaticais demonstram preferência pelo uso do sujeito nulo, tendo um total de 66% de nulos contra 34% de plenos:

Quadro 4 - Ocorrência de sujeitos nulos e plenos em PE, adaptado de Duarte (1995, p. 8)

Sujeito	Nulo		Pleno		Total	
Pessoa	Ocorrências	%	Ocorrências	%	Ocorrências	%
Primeira	334	60%	227	40%	561	100%
Segunda	101	73%	37	27%	138	100%
Terceira	303	73%	114	27%	417	100%
Total	738	66%	378	34%	1116	100%

Como podemos observar, é na primeira pessoa que as ocorrências de sujeitos plenos são mais frequentes (60%). Embora apresente uma desinência exclusiva, tanto no singular quanto no plural, ainda é a primeira pessoa que apresenta os índices mais baixos de sujeitos nulos. Conforme Duarte (1995, p. 9), essa não é uma exclusividade da língua portuguesa, há um estudo com a língua italiana, na sua modalidade oral, que aponta “o fato de que o falante geralmente não se apresenta numa conversa usando apenas o mecanismo de concordância verbo-sujeito (que viria a ser popularizado como característico das línguas pro-drop); na maioria dos casos, um pronome pleno é utilizado. Ainda sobre a primeira pessoa, é importante ressaltar o pouco uso da expressão *a gente* de referência definida, e enfatizar que em seu pouco uso, ainda houve falantes com nível de escolaridade mais baixo que produziram a forma verbal trazendo a desinência *-mos*.

A segunda pessoa, com 73% de ocorrência de sujeitos nulos, mantém um nível bem elevado da categoria vazia. Esse resultado corrobora a afirmação feita anteriormente de que o pronome *você* ainda é mantido como forma de tratamento no PE, assim não prejudicando a representação do sujeito.

2.3.2 O parâmetro do sujeito nulo no português brasileiro

Neste tópico, abordaremos duas propostas que se referem ao parâmetro do sujeito nulo no português brasileiro. Iniciaremos com a proposta de Duarte (1993, 1995) que afirma que essa língua está passando por uma mudança paramétrica, passando de língua pro-drop para língua não pro-drop. Em seguida, abordaremos a proposta de Rodrigues (2004), Holmberg

(2005) e Kato e Duarte (2014) que afirmam que o PB é uma língua de sujeito nulo parcial, sendo essa abordagem assumida por nós nesta dissertação.

2.3.2.1 A proposta de Duarte (1993, 1995)

Em seu estudo diacrônico, Duarte (1993) analisou textos de peças teatrais para averiguar os efeitos que a simplificação dos paradigmas flexionais estava causando, pois, nessa língua os pronomes *você* e *o(s) senhor(es)* também substituem as formas *tu* e *vós*, respectivamente. Diferentemente do PE, o PB usa as formas de terceira pessoa – singular e plural – como as únicas possibilidades de referenciar a segunda pessoa, ultrapassando-se, assim, o limite de sincretismos estabelecidos por Roberts (1993). Nesse caso, perde-se a regularidade na preferência pelo sujeito nulo. Ainda de acordo com Duarte (1993), a expressão *a gente* substitui o pronome *nós* nessa língua, situação que resulta em um paradigma com apenas três formas distintivas. A seguir apresentaremos três quadros que mostram a evolução sofrida por esses paradigmas até chegar às três formas distintivas. Vejamos:

Quadro 5 - Paradigma de flexão verbal do português brasileiro, adaptado de Duarte (1995, p. 32)

Português brasileiro	
1ª pessoa do singular	(Eu) canto
2ª pessoa do singular	(Você/Tu) canta(s)
3ª pessoa do singular	Você/ Ele canta
1ª pessoa do plural	(A gente/Nós) canta(mos)
2ª pessoa do plural	(Vocês/Vós) cantam/cantais
3ª pessoa do plural	Vocês/ Eles cantam

No quadro 5 vemos um paradigma flexional ainda com seis formas distintivas como as que são apresentadas no PE, porém os pronomes já apresentam espaço para o *Você* concorrendo com o *Tu* de segunda pessoa do singular e *Ele* de terceira pessoa do singular, quando observamos o plural, temos a expressão pronominal *A gente* concorrendo com o *Nós* de primeira pessoa e o *Vocês* concorrendo com *Vós* e *Eles* pronomes de segunda e terceira pessoa respectivamente. Vejamos o quadro 6 com a evolução do paradigma:

Quadro 6 - Paradigma de flexão verbal do português brasileiro, adaptado de Duarte (1995, p. 32)

Português brasileiro	
1ª pessoa do singular	(Eu) canto
2ª e 3ª pessoa do singular	(Você/Tu/Ele) canta
1ª pessoa do plural	(A gente/Nós) canta(mos)
2ª e 3ª pessoa do plural	(Vocês/Eles) cantam

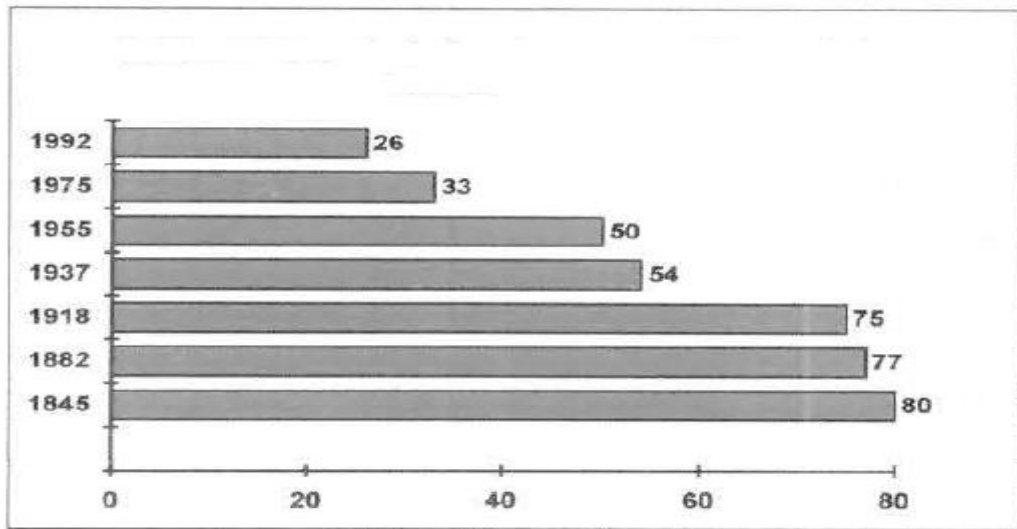
No quadro 6, podemos perceber o apagamento da flexão verbal *cantaS* em concordância com a 2ª pessoa do singular e o apagamento de *cantaIS* em relação à 2ª pessoa do plural. Nesse momento, o *Você* ainda concorre com o pronome de segunda pessoa do singular (*Tu*), porém já aparece substituindo o pronome de segunda pessoa do plural (*Vós*). Por fim, vejamos o quadro 7 com a última evolução e as três formas distintivas do paradigma flexional.

Quadro 7 - Paradigma de flexão verbal do português brasileiro, adaptado de Duarte (1995, p. 32)

Português brasileiro	
1ª pessoa do singular	(Eu) canto
2ª e 3ª pessoa do singular	Você/Tu/Ele canta
1ª pessoa do plural	A gente canta
2ª e 3ª pessoa do plural	Vocês/ Eles cantam

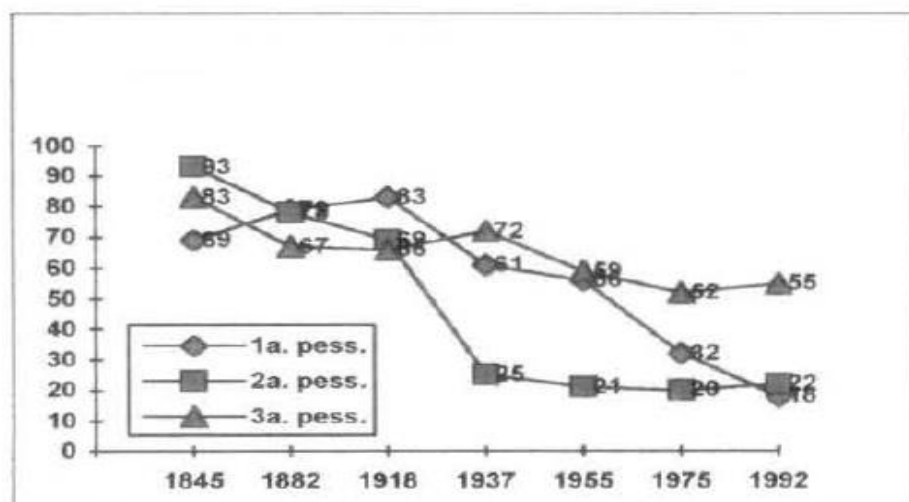
Podemos verificar, conforme o quadro 7, que o paradigma flexional passa a ter apenas três formas distintas: *cantO* – *cantA* – *cantAM*. Essa mudança acarretou diretamente na perda gradual da preferência pelo sujeito nulo no PB. No gráfico 1, podemos observar a ocorrência dos sujeitos nulos no PB, nos períodos analisados pela autora:

Gráfico 1 - Ocorrência de sujeito nulo em PB através de sete períodos. (%). Extraído de Duarte (1995, p. 19)



Nesse gráfico, podemos observar que, de 1845 a 1918, os índices de sujeitos nulos eram bem altos, à semelhança do que se tem observado para o PE. Nesse momento, seis formas distintas podem ser verificadas no paradigma pronominal, caracterizando-se a morfologia de flexão verbal como rica. No entanto, a partir 1937, a opção pelo sujeito nulo começa a diminuir, refletindo a perda da segunda pessoa direta. Por fim, por volta de 1975, a forma *a gente* é concorrente do pronome *nós*, causando uma drástica diminuição de sujeitos nulos e a preferência por sujeitos plenos no PB. Ainda, de acordo a autora, o processo de mudança do PSN não se dá da mesma maneira em todas as pessoas gramaticais. Podemos observar essa variação no gráfico 2:

Gráfico 2 - Ocorrência de sujeitos nulos nas três pessoas em PB. (%) Extraído de Duarte (1995, p. 20):

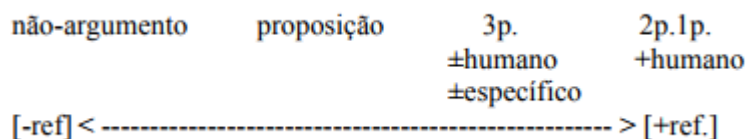


Conforme verificamos, na primeira pessoa, a ocorrência de sujeitos nulos tem uma queda extrema, passando de 69% no ano de 1845 para chegar a 18% em 1992. Na segunda pessoa, há uma variação de 93% para 22% dentro do mesmo período. Já a terceira pessoa, é a única que não apresenta uma variação extrema: a queda acontece de 83% para 55% apenas. Esses resultados colocam-nos diante de um fato: há preferência por sujeitos plenos de primeira e segunda pessoa, sendo mais produtivos sujeitos nulos de terceira pessoa. É importante ressaltar que, dentro dessa perspectiva, surge a ideia de que o PB é uma língua de orientação para o tópico, pois os sujeitos nulos de terceira pessoa são licenciados porque seus referentes já estão dados no domínio do discurso, tal como é defendido para o chinês, por exemplo.

Como visto, em sua tese, Duarte (1995) observou como a perda do princípio “Evite Pronome” estava acontecendo no PB. Por meio de uma amostra sincrônica da fala espontânea de universitários e a fala veiculada pela mídia, ela constatou que o PB ainda conta com uma ocorrência expressiva de sujeito nulo na terceira pessoa gramatical – fato que corrobora com os resultados de Duarte (1993) –, embora a preferência por sujeitos plenos seja nítida. E ainda o sujeito nulo também não é mais produtivo com referentes cujo traço é [-animado], principalmente, quando observada a fala dos mais jovens que preferem o sujeito pleno nesse contexto.

Vale referirmos que Cyrino, Duarte e Kato (2000), baseadas nas evidências de Duarte (1995), propõem uma hierarquia de referencialidade que vai atuar nos processos de mudança que envolvem os pronomes. Segundo elas, para uma língua como o PB, que tem a opção de variantes nulas e não-nulas, a hierarquia de referencialidade vai ser determinante na escolha de uma dessas variantes. No caso dos sujeitos plenos, os itens que são mais referenciais são os primeiros a serem plenos; é o que acontece com a primeira e a segunda pessoas – que por sua vez, apresentam um traço [+ humano] e estariam em um extremo à direita na hierarquia proposta. Já a terceira pessoa que apresenta traços [+/- humano], [+/- animado], [+/- específico], estaria mais à esquerda dessa hierarquia, sendo um ambiente menos propício à produção de sujeitos nulos. Vejamos a representação da hierarquia proposta pelas autoras que mostra um contínuo do [-referencial] ao [+referencial]:

Figura 1 - Hierarquia Referencial. Adaptado de Cyrino, Duarte e Kato (2000, p. 59)



Como podemos perceber, argumentos [+ humano] estão mais à direita e tendem a ser plenos, enquanto os não-argumentos à esquerda do contínuo tendem a ser nulos. Vejamos a seguir em (4) e (5):

(4) Eu estou doente.

(5) A capa do meu dicionário_i parece como se _itivesse sido rasgada.

Tomando por base os traços presentes na escala de referencialidade, vemos em (4) um argumento [+ humano] e que foi expresso na oração, um pronome de primeira pessoa, enquanto que em (5) há a terceira pessoa em um ponto mais baixo da escala de referencialidade, devido à interação de traços [- humano], apresentando um sujeito nulo.

Ademais, Duarte (1995), comparando o PB com o francês cuja morfologia de flexão verbal é fraca na atual sincronia, verificou que esta passou de uma língua pro-drop para uma língua não pro-drop em decorrência do enfraquecimento do seu sistema de flexão verbal ao longo dos séculos, uma situação que vem sendo observada na atual sincronia do PB.

Nesse sentido, a autora argumenta que o PB perdeu o princípio “Evite Pronome” e, como consequência, caminha na direção das línguas não pro-drop.

2.3.2.2 A proposta de Rodrigues (2004), Holmberg (2005) e Kato e Duarte (2014)

Contrariamente a Duarte (1995), estudos recentes consideram o PB uma língua de sujeito nulo, mas, devido ao fato de seus sujeitos nulos estarem submetidos a restrições não observadas em línguas de sujeito nulo consistente, como o PE e o italiano, trata-se de uma língua de sujeito nulo parcial, conforme Rodrigues (2004), Holmberg (2005) Kato e Duarte (2014). Listamos as seguintes restrições analisadas por Kato e Duarte (2014):

- a) Restrições semânticas de referencialidade: “verificou-se, no PB, que quanto mais referencial é o sujeito maior a expectativa de um pronome expresso” (KATO; DUARTE, 2014, p. 17).
- b) Restrições de propriedade morfossintáticas: “o PB respeita o EPP seja movendo um DP para a posição de Spec, TP ou movendo um elemento fraco para uma posição de adjunção a T”. (KATO, DUARTE, 2014, p. 17) O uso de sujeitos nulos em PB está restrito a alguns contextos frásicos.

Como vimos, Kato e Duarte (2014) apresentam contextos restritos para o uso de sujeito nulo em PB, alguns desses contextos também tinham sido previstos por Holmberg, Nayadu e Sheehan (2009). Estes apresentam o sujeito nulo em sentenças genéricas e em construções encaixadas complementos, enquanto aquelas preveem o nulo genérico, mas não apresentam o nulo em sentenças complemento, as autoras também fazem uma previsão para os expletivos nulos, o que as destaca dos autores citados acima, já que eles não fazem menção aos expletivos nulos. Abaixo descrevemos os contextos apresentados por Kato e Duarte (2014):

- a) O sujeito nulo em sentenças encaixadas – essas sentenças apresentam apenas uma leitura no PB, diferentemente de línguas de sujeito nulo prototípico.
- b) O sujeito nulo genérico – que se tornou possível com o desaparecimento do *se* indefinido.
- c) O sujeito nulo expletivo – são os sujeitos nulos mais baixo na hierarquia.

2.4 O PARÂMETRO DO SUJEITO NULO E A MORFOLOGIA DE FLEXÃO VERBAL NAS VARIEDADES AFRICANAS DO PORTUGUÊS

De acordo com os estudos já desenvolvidos, faremos aqui uma síntese de alguns dos resultados já alcançados nas variedades africanas do português, a fim de verificarmos se há semelhanças com o PE e/ou com o PB no que se refere ao PSN.

Contudo, antes dessa abordagem, acreditamos que é importante discorrermos um pouco sobre estudos que mostram o fato de as pessoas mais escolarizadas apresentarem um uso mais próximo ao da norma europeia do português (cf. Mota, 2015; Mota, Miguel e Mendes 2012). Dessa forma, temos em Brandão e Vieira (2012) que quanto mais escolarizado

é o informante, mais baixa é a percentagem de não marcação de concordância verbal tanto no PST quanto no PB urbano analisados pelas autoras, logo, vemos que a escolarização é o fator mais relevante em ambas as variedades. Na pesquisa realizada por Bazenga (2016), temos que o nível de escolarização superior é um fator determinante na concordância de p6, atingindo cerca de 90% de concordância nos níveis mais altos de escolarização, verificamos também que com o aumento do acesso à escola, pela faixa etária mais jovem, aumenta-se também as marcas explícitas de concordância. Esse fato é também apontado por Scherre & Naro (2006) e por Lucchesi (2006).

Como visto, entendemos que a variável escolaridade, embora não seja abarcada por esse trabalho, tem um papel importante nos resultados já obtidos em outras pesquisas, pois, vimos que quanto mais escolarizado é o indivíduo, mais tende a aproximar-se da norma culta do português. No caso das variedades africanas da língua portuguesa, quanto mais escolarizado, mais o indivíduo tende a aproximar-se do PE que é a língua ensinada nas escolas.

2.4.1 Angola

Oliveira e Santos (2007) realizaram um estudo com dados de fala para analisar a presença/ausência de sujeitos nulos pronominais na posição de sujeito no português angolano (PA) em comparação com o PB e o PE e concluíram que o PA é uma língua que, assim como o PE, tem preferência por sujeitos nulos, apesar de perceberem que, em relação à 1ª pessoa do singular e à 1ª pessoa do plural, o PA adéqua-se à hierarquia da referencialidade, pois, tal como o PB, os sujeitos tendem a ser plenos com essas pessoas, obtendo nos resultados cerca de 70% de sujeitos plenos nas primeiras pessoas do singular e do plural. Observemos no quadro 8 o número de ocorrências das formas pronominais na posição de sujeito em PA:

Quadro 8 - Número de ocorrências das formas pronominais na posição de sujeito em PA em 369 sentenças. Adaptado de Oliveira e Santos (2007, p. 97)

Pronome	Preenchido	Nulo
Eu	87	198
Tu	1	17
Você	1	140
Ele/Ela	10	11

Nós	38	59
A gente	1	-
Vocês	7	13
Eles/Elas	2	6

Como dito acima, os maiores índices de preenchimento do sujeito acontecem nas primeiras pessoas do singular e do plural, assim, mostrando uma afinidade com a hierarquia de referencialidade: “Quanto mais referencial, maior a possibilidade de um pronome não nulo” (Cyrino, Duarte & Kato, 2000, p. 59).

Teixeira (2008), em sua pesquisa, constata que o pronome *você* está sendo usado para tratamento íntimo ou universal na modalidade oral do PA, mais frequentemente, entre a população não escolarizada e por falantes nativos das línguas nacionais. Como vimos no quadro 1 deste capítulo, a língua materna falada por cerca de um terço dos angolanos é o umbundo, e o kimbundo vem, em segundo lugar, sendo falada pela quarta parte da população. Assim, entendemos que o convívio do PA com as línguas africanas é constante e que esse convívio pode interferir no uso dessa variedade do português.

Conforme Teixeira (2012), em Luanda, não se registra um número expressivo da expressão pronominal *a gente* mesmo entre falantes não-alfabetizados, e, por conta disso, a concordância não foi bastante enfraquecida. Porém, nesse segmento da população, registra-se a preferência pelo *você*, o que ocasiona uma perda da diferença entre a segunda e a terceira pessoa, indo na direção da gramática do kimbundo e distanciando-se da gramática do PE. Vejamos o quadro abaixo com o paradigma do quimbundo:

Quadro 9 - Paradigma flexional do kimbundo, adaptado de Teixeira (2012, p. 10)

Kimbundo	
Eme ngidya	Eu como
Eye udya	Tu come
Mwene udya	Ele come
Etu tudya	Nós comemos

Conforme o quadro 9, não há marcas de concordância entre o pronome sujeito e o verbo, esse fato diferencia o kimbundo do PE que, conforme o quadro 3, apresenta sistema flexional rico e marcas de concordância entre o pronome sujeito e o verbo.

Ainda nessa mesma pesquisa, a autora diz que, dos 434 dados de fala analisados, 65% dos resultados apresentaram o uso do sujeito expreso contra 35% de sujeitos nulos, indo na direção do que foi verificado por Oliveira e Santos (2007). A seguir destacamos duas sentenças que foram trabalhadas na pesquisa de Teixeira (2012), em (6) e (7) observamos o uso do pronome *você*, que diferentemente do PB, apresenta desinência verbal de segunda pessoa, quando usado por falantes de baixa ou nula escolaridade:

(6) Vai lá que você cresceste na cidade, mas se você tinhas crescido aqui no mato...(F2AK1 reportando uma conversa de seu pai com o futuro marido dela)

(7) Se tu não encostar, pode descer. (F2AP)

(8)

(TEIXEIRA, 2012, p. 10)

Como vemos em (6) há o uso da desinência verbal de segunda pessoa, porém em concordância com o pronome *você*, que no PB assume a desinência verbal de terceira pessoa. Enquanto que em (7) o pronome *tu* aparece expreso e nulo respectivamente, ambos com desinência verbal de terceira pessoa, o que já é frequente no PB.

O PA, portanto, ao que tudo indica, passa por processo de mudança rumo ao declínio de sujeitos nulos. Contudo, o fato de Angola ser um país, cuja independência política é recente, mantendo viva ainda a cultura de elite – nesse caso, a elite portuguesa – pode acarretar um freio a essa mudança. Sendo assim, ressaltamos que o PA ainda não tem uma identidade definida.

2.4.2 Moçambique

Segundo Bravin dos Santos (2010), o PM assemelha-se ao PE no que se refere ao PSN. Comparando essas variedades do português com o PB no tocante à acessibilidade do antecedente, verificou que há um comportamento diferente do sujeito anafórico de 3ª pessoa: enquanto no PB há preferência pela forma plena do sujeito anafórico de 3ª pessoa, o PE e o PM optam pelo sujeito nulo. Nas sentenças (8) e (9) abaixo podemos observar como essa variedade estudada no âmbito da terceira pessoa se comporta:

(9) ele_i regressava de morrumbene havia pressão ele_i saía __ (cv) ,mudava de local (PM)

(10) a pessoa_i não estudava ou então a pessoa_i convertia-se para a outra igreja (PM)

(BRAVIN DOS SANTOS, 2010, p. 30)

Como dito anteriormente, há uma diferença no comportamento do sujeito anafórico de terceira pessoa entre o PB, o PE e o PM. Enquanto aquele tem preferência pela forma plena, estes têm preferência pela forma nula do sujeito. Vejamos essa diferença no quadro abaixo:

Quadro 10: Ocorrência de sujeitos nulos em PM, PB e PE extraída de Bravin dos Santos (2010, p. 31)

PM			PB			PE		
Apl	%	Total	Apl	%	Total	Apl	%	Total
102	72	141	136	37	361	126	78	162

Ainda é possível observarmos que o contexto em que o índice de sujeito nulo é alto no PB é o mesmo em que sua realização é categórica (ou quase categórica) no PM e no PE. Isso explica porque a 3ª pessoa se constitui um elemento de resistência à mudança do parâmetro do sujeito nulo no PB e confirma a atuação de hierarquia de referencialidade (cf. CYRINO, DUARTE, KATO, 2000).

Oliveira (2016) utilizando-se das ferramentas metodológicas da sociolinguística, buscou elucidar as características que estariam relacionadas à expressão do sujeito na variedade moçambicana do português e obteve resultados que se distanciaram da pesquisa de Bravin dos Santos (2010). De acordo com o autor, parece estar ocorrendo um “comportamento misto” no que concerne às formas de 2ª e 3ª pessoa, as propriedades da posição de sujeito observadas no português moçambicano aparentam estar a meio caminho do PE e do PB:

O resultado da investigação permite entrever um quadro que põe o português falado em Maputo a “meio caminho” do português europeu e do português brasileiro: se não é possível afirmar que o português moçambicano tende, na sua gênese, a se afirmar como uma língua que mostra preferência pela realização fonológica do sujeito, igualmente, não é possível concluir que as propriedades da posição de sujeito observadas no português europeu permanecem preservadas na nova variedade. (OLIVEIRA, 2016, p. 10-11)

Diante do exposto, percebemos que existe a possibilidade de a variedade moçambicana do português estar também passando por um processo de mudança no que tange à marcação do parâmetro pro-drop, contudo, entendemos que para chegar a essa hipótese é necessário a realização de uma pesquisa diacrônica. Não há nessa pesquisa uma abordagem sobre o contato do PM com as línguas africanas faladas em Moçambique, no que concerne ao PSN. Por fim, entendemos a necessidade de mais pesquisas que abordem a comparação dessas variedades – africanas, brasileira e europeia –, com o intuito de termos resultados mais expressivos sobre esse possível processo de mudança.

2.4.3 Cabo Verde

Alexandre, Duarte e Santos (2013), analisando o crioulo cabo-verdiano, identificaram que esta é uma língua de sujeito nulo parcial, assim como o PB. As autoras afirmam que a probabilidade de ocorrer sujeitos nominativos em domínios não finitos é uma característica das línguas peninsulares, das quais o PB e o crioulo cabo-verdiano compartilham. O crioulo cabo-verdiano é uma língua de base lexical portuguesa com aspectos gramaticais de origem africana, nascida do contato linguístico entre o português médio e várias línguas africanas. Conforme Alexandre, Duarte e Santos (2013), há uma porcentagem muito baixa de palavras de origem africana no vocabulário nuclear do crioulo cabo-verdiano, 90% de palavras da língua portuguesa para 10% de palavras das línguas africanas, contudo, há grandes evidências de traços gramaticais das línguas africanas nesta língua. À exemplo, temos a expressão do tempo e do aspecto verbais que são semelhantes a processos encontrados em línguas africanas. Esses processos referem-se à oposição do tempo passado vs. o tempo presente ou para marcação de perfectividade (com verbos dinâmicos) vs. imperfectividade (com verbos estáticos), como podemos observar em (10) e (11):

- (11) a. *N ka Ø kume keju.*
 1SG NEG PFV comer queijo
 ‘Eu não comi o queijo.’
 b. *Mininus ta djuga karanbola.*
 meninos IPFV jogar carambola
 ‘Os meninos jogam ao berlinde.’

- (12) a. *E fase kel trabadju-li.*
 3SG fazer(PFV) DEM trabalho-PROX
 ‘Ele fez este trabalho.’

(ALEXANDRE, DUARTE e SANTOS, 2013, p. 9)

Conforme vimos nos exemplos, o sistema de expressão de tempo, modo e aspecto são diferentes entre a maioria das línguas africanas e as línguas ditas ocidentais, pois o sistema das línguas africanas recorre a morfemas livres que ocorrem antes do verbo. As autoras ainda afirmam que línguas crioulas, especificamente o crioulo cabo-verdiano, são línguas morfologicamente pobres, pois essas línguas não dispõem de morfemas próprios que expressem a ausência de tempo. Por fim, as autoras fazem uma análise em dados linguísticos de aquisição que confirmam as análises dos dados da gramática adulta do português brasileiro vernacular e chegam a conclusão, assim como outros autores chegaram, de que o PB está passando por um processo de enfraquecimento de flexão verbal. Essa conclusão é compartilhada por Galves (2001), Duarte (1995) entre outros. Alexandre, Duarte e Santos (2013) afirmam que mesmo com o enfraquecimento da flexão verbal, no caso do PB, e a pobreza de flexão verbal, no caso do crioulo cabo-verdiano, não implica no desaparecimento de uma propriedade característica de línguas de sujeito nulo, sejam elas consistente ou parcial, que seria a possibilidade de legitimar sujeitos nominativos em orações infinitivas. As autoras ratificam que essa é exatamente a propriedade que elas encontraram no crioulo cabo-verdiano e que consideram determinante para o estatuto dessa língua como língua de sujeito nulo parcial.

Lopes e Campos (2015) observaram que há no PCV uma tendência ao preenchimento da posição de sujeito das orações principais, principalmente quando se trata da 1ª pessoa do singular, em comparação com a 1ª pessoa do plural, como podemos observar em (12) e (13) extraído dos autores:

- (13) O problema que **eu** tenho com o português ou que **eu** tive durante o curso com o português, foi um problema que **eu** trouxe... **eu** carreguei do 12º, posso dizer do ensino secundário. Então [—] temos de falar de um ensino já, [—] temos de falar da qualidade já do sistema do ensino em Cabo Verde, e não do ensino superior.
 (PCV)

- (14) É precisamente isso. Então, onde é que [—] podemos ver as falhas no nosso sistema de ensino em Cabo Verde? Quais são essas falhas que nós temos que faz com que essa qualidade não seja assim uma coisa tão óbvia? (PCV)

(LOPES e CAMPOS, 2015, p. 234)

Os exemplos acima deixam claro a preferência pelo preenchimento da posição de sujeito na 1ª do singular em detrimento do preenchimento da posição de sujeito da 1ª pessoa do plural. Os autores acreditam que as relações de contato linguístico e os fatores sociolinguísticos que se verificam em Cabo-Verde são aspectos que influenciam a sintaxe do PCV, aproximando-o, em alguns casos, do PB. Lopes e Campos (2015, p. 326) afirmam que “Cabo Verde é uma nação diglósica (por questões de ordem política e não linguística), mas os cabo-verdianos são falantes bilíngues tendo o CCV como a língua materna e com diferentes graus de proficiências no PCV”. Para realização dessa pesquisa, os dados coletados a partir de alunos do nível superior (em formação ou já formado) que estudavam nas universidades em Cabo Verde, segundo os autores, neste nível acadêmico os informantes já possuem certa competência em língua portuguesa sem a influência direta do PE e do PB que recebem os alunos que se deslocam para prosseguir os estudos em Portugal e Brasil. Nosso estudo presa pelo contrário, buscamos compreender através dos alunos são-tomenses e guineenses que estão no Brasil como se dá a ocorrência do parâmetro do sujeito nulo nessas variedades africanas do português e perceber, através do tempo de estadia desses alunos no Brasil (os dados coletados nos primeiros e nos últimos períodos), se há influência do PB na frequência de sujeitos nulos e plenos.

2.4.4 São Tomé e Príncipe

Brandão e Vieira (2012), ao pesquisarem a concordância nominal e verbal no PB e no PST, verificaram que a regra de concordância é altamente produtiva entre os indivíduos mais escolarizados que vivem em áreas urbanas de São Tomé e que os índices percentuais para a aplicação da regra de concordância nominal e verbal são idênticos (93%). Esse resultado está intimamente associado ao fato de esses falantes tenderem a um uso mais próximo da norma europeia do português considerada a língua de prestígio. Abaixo podemos observar o quadro com os resultados obtidos:

Quadro 11 - Distribuição dos dados de concordância (marcação) e não concordância (não marcação) nominal de número e verbal de 3ª pessoa plural nas três variedades estudadas. Adaptado de Brandão e Vieira (2012, p. 1040).

Variedade	Concordância Nominal		Concordância Verbal	
	Marcação	Não Marcação	Marcação	Não Marcação
PB	91,1%	8,9%	78,1%	21,9%
PST	93,4%	6,6%	93,1%	6,9%
PE	99,9%	0,04%	98,9%	1,1%

Além da variável extralinguística *escolaridade*, outra variável analisada foi o sexo dos informantes. As autoras verificaram que as mulheres são-tomenses são mais conservadoras, ao contrário dos homens, por desfavorecerem o cancelamento da marca de número no sintagma nominal e no sintagma verbal, enquanto os homens o favorecem. No PB, por sua vez, essa distinção entre os sexos não é atestada.

Diante desses dados e sabendo que o PST é quase língua materna nessa ilha e que o crioulo nela falado é uma língua não-pro-drop, segundo Hagemijer (2009), supomos que o PST é uma língua pro-drop, mas que pode ter algumas restrições, como é o caso do PB, já que o santome, língua de maior contato com o PST, não admite sujeitos nulos, o que pode influenciar as ocorrências do parâmetro do sujeito nulo no português são-tomense. Ainda assim, reiteramos que, até o presente momento, não encontramos nenhum trabalho que abordasse de forma detalhada o PSN na variedade são-tomense do português.

Em uma pesquisa, não muito extensa, Oliveira e Silva (2018) analisaram a distribuição do sujeito nulo de maneira comparativa entre o PE e o PST. As autoras coletaram dados da língua falada de dezesseis são-tomenses, a partir de entrevistas informais e disponíveis no YouTube, usaram sexo e nível de escolaridade como variáveis sociais e analisaram a distribuição de sujeitos nulos e plenos em contextos declarativos finitos. Oliveira e Silva (2018) constataram que os resultados parecem corroborar com o que diz Hagemijer (2009) sobre a emersão de uma nova variedade de português, pois a diferença percentual entre sujeitos nulos e plenos encontrada pelas pesquisadoras era mínima. Contrariando às pesquisas com o PE, observada, por exemplo, por Duarte (1995, 2000). As autoras também atribuem esse índice quase equiparado entre sujeitos nulos e plenos ao contato linguístico existente entre o PST e o santomé. Abaixo, vemos um quadro demonstrativo dessa pesquisa:

Quadro 12 - Quantitativo geral e percentual de sujeitos plenos e nulos. Extraído de Oliveira e Silva (2018)

Total geral de dados	Sujeito Pleno	Sujeito Nulo
460 (100%)	205 (44,5%)	255 (55,5%)

Por fim, como podemos observar no quadro acima, o PST é uma língua pro-drop, mas que pode ter algumas restrições, e retomando o que foi afirmado pelas autoras, temos que a frequência geral dos dados aponta para uma pequena diferença entre plenos (44,5%) e nulos (55,5%), ao contrário do que se observa em PE cujo percentual de sujeitos nulos é bastante acentuado (Duarte, 1995, 2000).

2.4.5 Guiné Bissau

Como foi mencionado, até o presente momento, não encontramos uma pesquisa que analisasse o parâmetro do sujeito nulo no português guineense. Porém, o trabalho de Castro (2013), pesquisa qualitativa, discorre sobre o crioulo guineense, língua materna falada por 80% da população, em concorrência com o PGB que é falado apenas por 13% da população como língua segunda ou língua terceira (cf. quadro 1 deste capítulo). Nesse crioulo, os verbos não apresentam flexão: a forma verbal é a mesma para todas as pessoas, sendo os sujeitos obrigatoriamente plenos, o que a caracteriza como uma língua não-pro-drop. Vejamos o quadro abaixo com o paradigma de flexão verbal:

Quadro 13 - Paradigma de flexão verbal do crioulo guineense. Adaptado de Castro (2013, p. 102)

Paradigma de flexão verbal do crioulo guineense	
Ami korta	Eu cortei
Abo korta	Tu cortaste
El korta	Ele cortou
Anos korta	Nós cortamos
Abos korta	Vós cortastes
Elis korta	Eles cortaram

Observando o quadro 13, vemos que não há paradigma de flexão verbal para o crioulo guineense, língua falada pela maioria da população de Guiné Bissau. Sabemos que o contato linguístico do português guineense com o crioulo guineense e as demais línguas étnicas é frequente e que esse contato pode trazer mudanças para o PGB. Assim, diante dos dados e sabendo que, quanto maior é o uso de outra língua distinta do português, maior é o afastamento da variedade africana do português em relação à variedade europeia, deduzimos que, dentre as cinco variedades africanas que aqui foram abordadas, o PGB é o que mais se distancia do PE.

2.5 OBJETIVOS

Geral:

Investigar a interferência do português brasileiro na língua falada e escrita de alunos guineenses e santomenses dos primeiros e últimos períodos da UNILAB-CE no que se refere à distribuição de sujeitos nulos e plenos.

Específicos:

- (a) compreender a história de contato linguístico em Guiné Bissau e São Tomé, tendo em mente a aquisição do português e sua relação com o PSN;
- (b) observar se os contextos de sujeitos plenos e nulos no PGB e no PST estão submetidos às mesmas restrições já verificadas no PB;
- (c) analisar a natureza da morfologia de flexão verbal nos dados de fala e de escrita em PGB e PST e suas possíveis implicações na distribuição de sujeitos nulos;
- (d) comparar o uso de sujeitos nulos e plenos no PGB e no PST, tomando por base as variáveis extralinguísticas;
- (e) verificar se há uma possível assimetria entre a língua falada e a língua escrita de estudantes guineenses e são-tomenses da UNILAB, tendo em mente que a primeira pode apresentar maior ocorrência de plenos do que de nulos.

2.6 HIPÓTESES

Geral:

Alunos dos últimos períodos da UNILAB prestes a se formarem podem revelar um uso que se diferencia, em alguns aspectos, da norma europeia do português em virtude do contato intenso com o PB: sujeitos plenos podem ser mais produtivos.

Específicas:

- (a) Estudantes guineenses, desde os primeiros períodos, tendem a produzir mais sujeitos plenos em decorrência do uso frequente de sua língua materna, o crioulo guineense, uma língua não pro-drop, ao contrário do que pode ser verificado em São Tomé, em que a maioria da população já tem o português, uma língua de sujeito nulo consistente, como primeira língua (HAGEMEIJER, 2009);
- (b) As restrições encontradas em PB para a distribuição de sujeitos nulos podem ser evidenciadas na língua falada e escrita dos estudantes dos últimos períodos de diferentes cursos da UNILAB em virtude do intenso contato do PGB e do PST com o PB;
- (c) O PST apresenta uma morfologia de flexão verbal rica, enquanto que o PGB apresenta mudanças em sua morfologia de flexão verbal e possíveis perdas de paradigmas. Esperamos encontrar essa disparidade entre as línguas logo nos dados dos primeiros períodos, nos quais, não haverá influência do PB. A riqueza ou não dessa morfologia flexional pode influenciar diretamente nas ocorrências de sujeitos nulos, quanto mais paradigmas forem encontrados, maior o número de ocorrências de sujeitos nulos.
- (d) O sexo feminino pode ser mais conservador que o sexo masculino no PGB e no PST: as mulheres podem produzir mais concordância verbal e, conseqüentemente, mais sujeitos nulos, embora todos estejam dentro da mesma faixa etária e exerçam a mesma função social: estudantes.

3 FUNDAMENTAÇÃO TEÓRICA

Neste capítulo discorreremos sobre os pressupostos básicos da teoria gerativa e centraremos nossa atenção no Modelo de Princípios e Parâmetros (CHOMSKY, 1981, 1986 e seguintes). Abordaremos também a Aquisição da Linguagem à luz dessa teoria, tendo em vista que o português não é adquirido da mesma maneira em São Tomé e Guiné Bissau: enquanto na primeira, o português é a primeira língua (L1) da maioria das pessoas (cf. HAGEMEIJER, 2007); na segunda, trata-se de uma segunda língua (L2), pois sua aquisição ocorre em geral via escolarização.

3.1 PRESSUPOSTOS BÁSICOS DA TEORIA GERATIVA

3.1.1 Fundamento filosófico

Até a primeira metade do século XX, estava em vigor a teoria defendida pelos behavioristas, segundo a qual os comportamentos humanos eram gerados externamente por meio de associações entre estímulo e resposta, sendo essas associações criadas por meio de repetições. Nesse sentido, o ser humano aprende a linguagem por meio de imitação. É a partir de 1957, com a publicação do livro *Syntactic Structures*, que Noam Chomsky, contrário ao pensamento behaviorista, propõe a Teoria Gerativa cujas bases causam uma revolução no contexto linguístico e das ciências da cognição da época. A partir dessa teoria, dá-se enfoque à natureza mental da linguagem humana, entendida como um órgão biológico que se localiza na mente-cérebro do ser humano.

Ao propor sua teoria, Chomsky afirmava que a tarefa da linguística era de explicar o caráter gerativo da língua, isto é, a possibilidade de gerar uma quantidade infinita de frases com uma quantidade finita de fonemas, morfemas, palavras e regras gramaticais, o que vai de encontro à perspectiva estruturalista segundo a qual é tarefa da linguística descrever um dado sistema linguístico, tomando por base as relações entre os elementos desse sistema.

Indo de encontro também aos estudos comportamentais, Chomsky afirmava que o ser humano tem uma capacidade inata para adquirir qualquer língua natural e que o estímulo linguístico externo era necessário, porém não suficiente para a aquisição de tal língua, ou seja,

havia uma capacidade inata do ser humano para falar, contrariando o que acreditavam os comportamentalistas. Essa insuficiência do estímulo linguístico externo foi exemplificada como o “Problema de Platão”. Platão argumenta que existem conhecimentos que não podem ter sido adquiridos por experiência, em um de seus diálogos, Sócrates tenta mostrar que um escravo pode ter conhecimentos sofisticados de geometria, mesmo sem nunca ter tido instruções formais sobre aquela matéria. Ao relacionarmos o problema de Platão com a aquisição de uma língua natural, temos que uma criança conhece princípios que regem sua língua, mesmo sem ter tido aulas sobre aquela língua. Assim, compreendemos que o input é desorganizado e que não contém todas as informações necessárias para a criança.

Em linhas gerais, a teoria gerativa busca entender, a partir de uma perspectiva mentalista, o que é e como o conhecimento linguístico é adquirido pelo falante e como é usado esse conhecimento.

Portanto, ao contrário do pensamento empirista que embasa as teorias estruturalista e comportamentalista, a teoria gerativa tem como base filosófica o racionalismo. Tal como proposto por René Descarte de que existe a faculdade de pensar inata a todo o ser humano; para Chomsky, a Faculdade da Linguagem também lhe é inata.

3.1.2 Definição de linguagem

Sabemos que alguns animais têm um sistema sofisticado de comunicação, como é o caso das abelhas, dos golfinhos, entre outros. Contudo, embora possuam um sistema de comunicação eficaz, podemos dizer que essa comunicação é limitada, pois é restrita a alguns sinais e sons a que estão habituados a fazer.

A capacidade da linguagem de criar e entender sentenças nunca antes ouvidas é exclusiva do ser humano. Apenas o homem consegue, independentemente do seu grau de escolaridade, criar frases infinitas a partir de um conjunto limitado de elementos linguístico, o que revela uma propriedade exclusiva da linguagem humana: a criatividade. Segundo Chomsky (1957, p. 13 apud LYONS, 1987, p. 5), a língua(gem) é entendida “como um conjunto (finito ou infinito) de sentenças, cada uma finita em comprimento e construída a partir de um conjunto finito de elementos”, ou seja, a linguagem está associada a capacidade humana de gerar infinitas frases com recursos limitados. É importante ter em mente também

que não basta apenas a produção (a forma fonética), mas também a compreensão das estruturas (a forma lógica). Assim, as estruturas frásicas, para serem gramaticais (boas) em uma língua, devem estabelecer interface com o sistema articulatório-perceptual (forma fonética) e com o sistema conceptual-intencional (forma lógica).

Por ser uma capacidade humana, entende-se que todo ser humano, desprovido de patologias, é capaz de adquirir qualquer língua natural e que as línguas apresentam propriedades comuns. Embora existam várias línguas no mundo, que apresentem diferenças entre si, é possível dizer que há semelhanças linguísticas entre elas, já que todas estão previstas a partir da mesma base biológica. Assim, entende-se que há características universais comuns a todas as línguas.

Enfim, conforme a Teoria Gerativa, a linguagem é entendida como um objeto biológico, é uma capacidade inata, faz parte do DNA humano, o que explica a incapacidade dos outros animais adquirirem tal linguagem. Segundo Luguinho, Resenes e Negrão (2012), a capacidade da linguagem é própria do ser humano e essa capacidade dá-se devido a uma Faculdade da Linguagem (doravante FL) – componente da mente que é restrito a espécie humana. Disso decorre que a aquisição de uma língua particular não depende de uma variável social ou geográfica, haja vista que a criança precisa apenas ser exposta aos dados linguísticos primários (*input*) a fim de que a Faculdade da Linguagem que lhe é inata seja “acionada”.

3.1.3 Sobre a Faculdade da Linguagem

Por definição, segundo Chomsky (1986, p. 22), a FL é “uma componente particular da mente humana”, ou seja, trata-se de um órgão presente no cérebro responsável pela linguagem (produção e compreensão de estruturas frásicas). Fazendo uma analogia, sabemos que o corpo humano é composto por diversos órgãos e que cada um deles desempenha uma função específica dentro de cada sistema; assim, na mente/cérebro, há uma “parte” específica que lida com a linguagem. Para Chomsky (1986, p. 22-23),

A natureza desta faculdade é o tema básico de uma teoria geral acerca da estrutura da linguagem, que tem como objetivo a descoberta do conjunto de princípios e de elementos comuns às línguas humanas possíveis; atualmente esta teoria é muitas vezes chamada ‘gramática universal’ (GU). [...] Pode-se encarar esta faculdade como um ‘mecanismo de aquisição da linguagem’, uma componente inata da mente humana que origina uma língua particular pela interação com a experiência vivida, ou ainda como um mecanismo que

converte a experiência num sistema de conhecimento atingido: conhecimento de uma ou outra língua.

Assim, compreendemos que a FL é o órgão que permite o desenvolvimento das capacidades linguísticas, funcionando como ‘uma espécie de planta arquitetônica biológica’ que proporcionará ao indivíduo adquirir um sistema de conhecimentos linguísticos complexo e organizado que vai além do *input* recebido inicialmente.

Chomsky (1957; 1971) ainda propõe que há um isolamento funcional dos processos algorítmicos da estrutura mental particular da linguagem em relação aos outros processos mentais, ou seja, a linguagem é um módulo único e específico da cognição humana.

Para a Teoria Gerativa, a mente humana é modular, possuindo diferentes componentes que são responsáveis por diferentes capacidades cognitivas. Esse conceito de ‘módulo’ é derivado do conceito de modularidade da mente, seguindo o pensamento de Fodor (1983), e opõe-se ao conceito de uniformidade da mente, segundo a qual a inteligência humana é um todo indivisível, capaz de dar contas de todas as tarefas cognitivas realizadas pelo homem. Contrária a essa visão, a Teoria Gerativa defende que a mente é composta por módulos que possuem propriedades específicas. Conforme Raposo (1992, p. 30):

A concepção modular da mente humana defende que esta é formada por módulos autônomos, cada um deles caracterizados por princípios e representações específicas. Estes módulos <<comunicam>> entre si em pontos determinados, mantendo uma interação complexa que determina as propriedades dos fenômenos mentais humanos. Assim, para além do módulo linguístico, a mente humana possui por hipótese um módulo matemático, um módulo musical, um módulo espacial (que lhe permite compreender mapas, por exemplo), um módulo de formação de conceitos, entre outros cujo descobrimento e caracterização pertence às ciências cognitivas.

Assim, compreendemos que a linguagem representa um dos módulos cognitivos da mente humana, possui sua própria natureza e funcionamento e é independente dos outros módulos da mente, lidando com a aquisição do conhecimento linguístico, produção e compreensão de estruturas linguísticas.

Vale referirmos que, dentro da FL, há módulos diferenciados para lidar com diferentes tipos de informações linguísticas que funcionam de maneira integrada e interdependente em um processo de interação e dinamicidade. Conforme Miotto, Silva e Lopes (2013, p. 20) “o

módulo que lida, por exemplo, com a determinação da referência para os pronomes é diferente do módulo que lida com a estruturação das sentenças das línguas”. A partir dos diversos módulos que integram a FL, a mente humana é capaz de adquirir e processar um sistema complexo e sofisticado como uma língua natural.

3.1.4 Sobre a Gramática Universal e o Modelo de Princípios e Parâmetros

Assumimos aqui com Chomsky (1957, 1986) que a FL é inata ao ser humano, modular, e a aquisição de uma língua particular ocorre porque nessa faculdade existe um dispositivo de aquisição da linguagem (também chamado de Gramática Universal, GU) que se refere ao estágio inicial em que estão presentes as propriedades comuns a todas as línguas (os Princípios), bem como as propriedades que variam entre as línguas (os Parâmetros). A GU corresponde, portanto, à fase pré-linguística da criança e, somente após sua exposição ao *input*, esse dispositivo filtra os dados linguísticos relevantes e proporciona a aquisição do conhecimento linguístico, partindo do estágio inicial ao estágio final (conhecimento linguístico atingido). Para tanto, valores paramétricos são fixados no processo de aquisição da linguagem e a gramática de uma dada língua particular é adquirida, conforme é defendido no Modelo de Princípios e Parâmetros proposto por Chomsky (1981, 1986 e seguintes).

Centrando nossa atenção nos Princípios, Chomsky (1981, 1986) argumenta que são comuns a qualquer língua natural, rígidos, invariáveis e determinam as semelhanças entre as línguas. O número de Princípios é finito e, dentre eles, destacaremos o Princípio de Projeção Estendido (EPP, do inglês Extended Projection Principle), segundo o qual todas as orações finitas possuem a posição de sujeito projetada nas línguas naturais. Observemos as frases (14) e (15) em português e inglês, respectivamente:

(14) a. Nós comemos o bolo.

b. __ Comemos o bolo.

(15) a. We eat the cake.

b. *__ Eat the cake¹³.

¹³ Essa frase só seria gramatical no contexto da frase imperativa: “Coma o bolo”. Contextos de imperativos constituem ambientes obrigatórios de sujeitos nulos nas línguas naturais. Por isso, não é levado para a análise quando se trata do Parâmetro do Sujeito Nulo.

Nas frases acima, temos duas situações: a) em (14a), há um sujeito realizado foneticamente (sujeito pleno), ao contrário de (14b) que possui um sujeito não realizado foneticamente (sujeito nulo), e b) (15a) possui obrigatoriamente a posição de sujeito preenchida por um sujeito pleno, sendo agramatical a frase em (15b) com o sujeito nulo. Com esses exemplos, percebemos a universalidade do Princípio: todas as línguas naturais projetam, de fato, a posição de sujeito. Contudo, existe uma diferença entre as línguas: o português possui sujeitos nulos, ao passo que o inglês, não. Essa distinção só poderá ser explicada pela noção de Parâmetro.

Conforme Chomsky (1981, 1986), os Parâmetros da GU são compostos de valores binários (+) e (-) e são o *locus* da variação interlinguística. Raposo (1992, p. 55) afirma que os Parâmetros são “uma espécie de comutadores linguísticos”, pois esses servem como interruptores que são ligados ou desligados conforme as informações linguísticas que as crianças recebem do ambiente linguístico. Por volta dos 5 anos de idade, as crianças já têm fixado os valores paramétricos de sua língua. A esse respeito, Raposo (1992, p. 55) afirma:

Quando todos os parâmetros estão ligados, a criança adquire uma <<gramática nuclear>> (em inglês, <<Core Grammar>>), isto é, um sistema complexo de conexões entre os princípios universais rígidos e os parâmetros, o qual determina de um modo altamente específico as propriedades de cada língua particular. A aquisição é assim completamente identificada com o crescimento e a maturação de UG, que passa de um estado apenas parcialmente especificado (com parâmetros por fixar) a um estado completamente especificado (com os parâmetros fixados), funcionando então como um sistema computacional.

Retomando às frases (14) e (15), percebemos a existência de um parâmetro: o Parâmetro do Sujeito Nulo (PSN). Quando o valor desse parâmetro é marcado positivamente, dizemos que a língua é de sujeito nulo (língua pro-drop) (ex.: português), quando, negativamente, a língua é de sujeito não-nulo (língua não-pro-drop) (ex.: inglês). A seguir, vejamos mais algumas frases do português e do inglês, respectivamente:

(16) __ Chove!

(17) a. *__ Rains!

b. It rains!

Conforme Xavier (2006), em português, por ser uma língua [+ sujeito nulo], sujeitos nulos ocorrem quando são referenciais ou expletivos, enquanto, em inglês, sujeitos plenos devem ocorrer nesses contextos, caso contrário, as frases tornam-se agramaticais.

À semelhança do inglês, no francês os sujeitos possuem o mesmo comportamento por ser uma língua não-pro-drop:

(18) It rained yesterday.

(19) Il a plu hier.

Os pronomes *It* e *Il*, em inglês e em francês, respectivamente, são não-referenciais e ocupam a posição do sujeito. Portanto, são os sujeitos sintáticos em (18) e (19), evidenciando assim o Princípio de Projeção Estendida.

3.1.4.1 O Parâmetro do Sujeito Nulo

É, portanto, a partir do Modelo de Princípios e Parâmetros que o PSN é proposto, haja vista que a realização fonética do sujeito é obrigatória em algumas línguas (ex.: inglês e francês) e não em outras (ex.: português e italiano). Chomsky (1981, 1986) e Rizzi (1982, 1986) argumentam que a riqueza da morfologia de flexão verbal (*Agr*, do inglês Agreement) é a responsável pelo licenciamento e identificação de sujeitos nulos em línguas de sujeito nulo (ex.: italiano), ao contrário do que se observa na natureza da morfologia em línguas não-pro-drop (ex.: inglês). Observem-se os seguintes paradigmas que evidenciam essa diferença:

Quadro 14 - Paradigmas de flexão verbal do italiano. Elaboração própria.

Paradigmas de flexão verbal do italiano	
Io	Parlo
Tu	Parli
Lui/Lei	Parla
Noi	Parliamo

Voi	Parlate
Loro	Parlano

Quadro 15 - Paradigmas de flexão verbal do inglês. Elaboração própria.

Paradigmas de flexão verbal do inglês	
I	Speak
You	Speak
He/She/It	Speaks
We	Speak
You	Speak
They	Speak

No quadro 14, podemos observar os seis paradigmas de flexão verbal distintivos do italiano, com a mudança que se dá em todos os morfemas número-pessoais. Enquanto que no quadro 15, há apenas duas formas distintivas dos paradigmas de flexão verbal, havendo mudança apenas no morfema número-pessoal de terceira pessoa do singular.

Analisando o PSN, Rizzi (1982, 1986) defende que esse parâmetro está associado a um conjunto de propriedades, a saber:

- a) Possibilidade de sujeitos nulos referenciais definidos em sentenças finitas:

(20) Parli portoghese bravamente [italiano]

(21) Hablas portugués muy bien [espanhol]

- b) Inversão livre do sujeito:

(22) Ha telefonato Carlo a il suo zio [italiano]

(23) Llamó el Carlos a su padre [espanhol]

c) Ausência do efeito *that*-trace

(24) Chii crede che ___ partirà? [italiano]

(25) ¿Quién usted cree que partirá? [espanhol]

d) Verbos finitos com morfologia flexional rica

(26) Ho cantato tutta la notte [italiano]

(27) He cantado toda la noche [espanhol]

(RIZZI, 1986, p. 273)

Línguas não *pro-drop*, como o francês e o inglês, não possuem essas propriedades, pois não se encontra sujeitos nulos nesses contextos, ou seja, a categoria vazia denominada *pro*, um elemento [+ pronominal] e [- anafórico].

Analisando uma língua como o chinês que não possui morfologia de flexão verbal rica, mas que possui sujeitos nulos, Huang (1984) argumenta que o licenciamento desses sujeitos ocorre sob uma restrição: estão ligados a um tópico discursivo. Ou seja, *pro* é licenciado quando seu referente já se encontra no domínio discursivo.

Ademais, Jaeggli e Safir (1989), analisando o PSN e a tipologia dos sistemas de concordância que licenciam os sujeitos nulos, defendem que o licenciamento de sujeitos nulos não está diretamente ligado a uma flexão verbal rica, mas sim, a uma ‘uniformidade morfológica’ dos paradigmas verbais das línguas. Essa uniformidade pode ser constituída apenas de formas ‘derivadas’ ou ‘não-derivadas’. Contudo, essa condição de uniformidade morfológica não consegue explicar amplamente o comportamento de todas as línguas, deixando de fora algumas línguas como o alemão e o sueco.

Ainda em relação à riqueza de *Agr* e ao PSN, Roberts (1993) propõe dois tipos de riqueza para caracterizar a morfologia de concordância verbal: a formal e a funcional. Para ele, verifica-se a riqueza formal de um paradigma quando há morfemas número-pessoais que identificam cada pessoa do paradigma, enquanto a funcional caracteriza-se pela existência de um sincretismo e de uma forma zero no paradigma. Assim, teríamos a riqueza formal para o paradigma flexional do PE e a riqueza funcional para o paradigma flexional do PB.

Galves (2001) formula uma análise alternativa para explicar a riqueza de Agr ao propor a existência de duas pessoas na GU: pessoa sintática e pessoa semântica, a primeira sintática é resultado da combinação entre traços binários relativos à pessoa e ao número, tal como é verificado por ela para o PB, ao passo que a segunda está relacionada à existência de morfemas que identificam gramaticalmente as três pessoas do discurso no singular e/ou no plural, a exemplo do que ocorre no PE.

A partir do exposto, assumimos aqui que (a) o PSN é composto de dois valores e que a escolha por um deles no processo de aquisição decorre da exposição da criança ao *input* de sua língua e (b) quatro sistemas de sujeitos nulos podem ser verificados, tomando por base a proposta de Roberts, Nayudu e Sheehan (2009), sendo o PE, a língua do colonizador em terras africanas e brasileiras, uma língua de sujeito nulo consistente devido à sua morfologia de flexão verbal que é rica, não estando os sujeitos nulos submetidos às restrições encontradas no PB, por exemplo.

Tendo como objetivo analisar comparativamente o PSN em São Tomé e em Guiné Bissau, não podemos esquecer que o português é adquirido de forma diferente nessas ilhas: enquanto na primeira, trata-se quase majoritariamente de uma primeira língua (L1), na segunda, adquire-se em geral como segunda língua via escolarização. Esse fato leva-nos a pensar que a variedade são-tomense do português tende a ser muito próxima da variedade europeia no que se refere à distribuição de sujeitos nulos, ao contrário da variedade guineense do português que pode apresentar mais sujeitos plenos em virtude da grande influência da língua materna, o crioulo guineense, uma língua não-pro-drop.

Dessa forma, interessa-nos saber em que sistema de línguas de sujeito nulos situam-se o PST e o PGB, levando em consideração seu processo de aquisição linguística, seu contato com outras línguas, e também, no caso da nossa amostra, a possível interferência do PB nessas duas variedades, sobretudo, quando observados os informantes africanos de nossa pesquisa (alunos da UNILAB) que estão nos últimos períodos de seus cursos e que, portanto, já possuem um intenso contato linguístico com essa variedade do português.

3.2 AQUISIÇÃO DE L1 E L2 À LUZ DA TEORIA GERATIVA

3.2.1 Aquisição de L1 à luz da Teoria Gerativa

A Teoria Gerativa entende que o processo de aquisição da linguagem é universal no sentido de que uma criança, quando exposta a uma determinada língua, tem condições de adquiri-la sem instruções e repetições de estruturas devido à FL que está em sua mente-cérebro e na qual encontra-se a GU. Nesse sentido, não existe uma língua mais difícil de se adquirir que outra e, por volta dos cinco anos de idade, a criança já adquiriu a gramática nuclear de sua língua.

Na década de 60, Chomsky rebate o Behaviorismo – teoria comportamentalista – afirmando que os dados linguísticos primários (a saber: o *input*) são insuficientes para a aquisição da linguagem. Esses dados externos são insuficientes, tendo por base o argumento defendido por esse linguista no que se refere à “pobreza de estímulo”, pois o *input* é incompleto, desorganizado, pobre – por não conter toda informação necessária – e não fornece informações sobre o que não é permitido na língua (sentenças agramaticais), pois apenas é composto de sentenças gramaticais.

Assegurando que os dados externos não são suficientes, Chomsky (1965) postula que existe uma estrutura inata codificada na Faculdade da Linguagem, que como já mencionamos, tem a ver com a GU. Miotto, Silva e Lopes (2013, p. 31) afirmam que

O processo de aquisição da linguagem, então, é tido como a ‘formatação’ da Faculdade da Linguagem através da fixação dos valores dos parâmetros previstos na UG. Como dissemos anteriormente, a UG é, nesse sentido, um quadro do estágio inicial da aquisição (conhecido como S_0) e o seu produto seria o estágio final da aquisição, isto é, o estágio em que a criança atinge a gramática adulta de sua língua ($S_{\square}$) (do inglês *stable stage*). Em termos linguísticos é bastante complicado falar em produto ou estágio final do conhecimento. Assim é mais plausível admitir-se que a gramática atinja um estágio de estabilização que seria considerado, então, como o estágio em que a criança apresenta uma gramática próxima à dos adultos ao seu redor.

Portanto, a GU é o estágio inicial do processo de aquisição da linguagem, sendo composta pelos Princípios e Parâmetros cujo valor deverá ser fixado pela criança durante sua

exposição ao input. Após a marcação dos valores paramétricas, a criança adquire a gramática da língua-alvo a que foi exposta, a que chamamos de estágio final da aquisição.

3.2.1.1 Aquisição de português como L1 em São Tomé e Príncipe

Como é sabido, a colonização de São Tomé ocorreu por volta de 1493 e escravos de vários países da África foram trazidos a essa ilha para trabalhar no cultivo da cana-de-açúcar. Segundo Hagemeijer (2009), do contato das diferentes línguas desses escravos com o português falado pelos senhores portugueses surgiu um *pidgin*, que, por sua vez, deu lugar a crioulos que aos poucos se tornaram estáveis, sistemáticos e estruturados, formando-se, assim, um contexto étnico e linguístico complexo. Lorenzino (1996) afirma que essa realidade linguística e social deu lugar a três crioulos que permanecem atualmente em São Tomé e Príncipe: o santomé ou forro que é a língua de São Tomé; o angolar que é a língua dos Angolares e o lung'ie que é a língua de Príncipe.

Pesquisas linguísticas apontam para o fato de que o santomé é a língua com maior número de falantes e que esse dá-se devido ao fato de ser a língua dos mestiços que atingiram um alto *status* econômico quando passaram a ser donos de terra e de escravos. O lung'ie, por sua vez, não sofreu grande interferência de grupos influentes como a igreja, os portugueses e os donos de terra (mestiços). Por fim, pouco se sabe do angolar, contudo, acredita-se que sua origem esteja ligada à fuga de escravos que se refugiaram e procuraram um estilo de vida mais centralizado e recluso, não sofrendo muita interferência de grupos sociais influentes, conforme Lorenzino (1996).

Diante do exposto, a constituição linguística de São Tomé e Príncipe vive em constante multilinguismo, tendo o português um papel fundamental nas esferas política e administrativa, dividindo o espaço com as línguas crioulas no cotidiano da sociedade. Nessa ilha, a língua do colonizador é a língua oficial da nação e tem prestígio em relação às outras línguas que convivem no país. É a língua da escolarização, do ensino, do meio político, administrativo e da comunicação social e midiática.

Vale referimos que o português fazia parte de uma minoria no início da colonização, mas foi ganhando muito prestígio entre os falantes são-tomenses ao ponto de se ser falada por mais de 90%, vindo a se constituir, segundo Hagemeijer (2009) e Gonçalves (2010), a

primeira língua (a língua materna) da maioria desses falantes. A pesquisa de campo por nós realizada evidencia, por exemplo, que todos os alunos são-tomenses afirmaram que o português é a sua língua materna. Além disso, é importante ressaltar que o *input* disponível para a formação do português são-tomense como língua materna é proveniente de variedades interlinguísticas (o português vem passando por um processo de mudança na aquisição, deixando de ser segunda língua e tornando-se língua materna em quase todo território) do português como segunda língua em contato com línguas crioulas. Sobre o processo de transição do português como segunda língua para língua materna em São Tomé, Gonçalves (2010, p. 20) conclui o seguinte: “[o]s elementos estatísticos de que dispomos levam-nos a considerar que o processo de transmissão do português L2 à mais jovem camada da população, competindo com as línguas nativas e afirmando-se como L1 perante elas, impulsiona a emergência de uma nova variedade de língua.”

Esse processo de transição de uma L2 para uma L1 também passa pela perda da competência linguística da língua materna (nesse caso, o santomé) que é motivada pela preferência do uso da língua “nova” (o português) em detrimento daquela. Essa perda tem sido descrita como um fenômeno complexo chamado de erosão linguística. Conforme Gonçalves (2010, p. 20-21):

a erosão linguística envolve cinco estádios, sendo que o primeiro contempla a L1 enquanto único sistema linguístico de que o falante dispõe. O segundo estádio corresponde à fase em que a L1 concorre com a L2, num período de transição de monolinguismo a bilinguismo. Neste período, o uso da L1 é praticamente restrito a ambientes informais, tipicamente familiares, enquanto a L2 é a língua oficial, de prestígio, usada na comunicação social, no ensino, na administração e noutros contextos formais. No terceiro estádio, numa permanente diglossia, os falantes adoptam maioritariamente a L2 em detrimento da primeira língua, apontando para o seu apagamento e perda. Por fim, no quarto e quinto estádios, respectivamente, a proficiência na L1 pela comunidade de falantes é cada vez mais limitada, culminando na completa “extinção” da L1 e consequente substituição pela L2. A delimitação dos cinco estádios não é claramente visível, uma vez que a erosão linguística opera em tempos distintos nos vários falantes.

Diante do exposto, assumiremos para este trabalho, seguindo Gonçalves (2010), que a situação linguística em São Tomé vivencia os três últimos estádios, pois o uso das línguas crioulas – principalmente o santomé – vem sendo restrito a situações informais, porém não podemos afirmar que esteja caminhando para o desaparecimento dessas línguas, uma vez que, conforme vimos no quadro 1, ainda temos mais de 70% da população usando o santomé. Esse

índice alto de falantes do santome pode ser relacionado com a faixa etária, a escolaridade e o meio (rural ou urbano) em que esse falante vive. Tomando por base que em São Tomé o português é uma L1, esperamos encontrar resultados semelhantes ao do PE, no que concerne ao uso do sujeito nulo.

3.2.2 Aquisição de L2 à luz da teoria gerativa

Conforme Leffa (1988), ao aprender uma língua estrangeira, a criança é exposta ao *input* em um contexto totalmente artificial, de maneira estruturada e restrita, como acontece no contexto escolar. Por sua vez, a aquisição de uma segunda língua não se encontra restrita ao universo escolar, pois ocorre também nas diversas esferas sociais do cotidiano, tais como: o contexto midiático, administrativo e outros.

Em se tratando do contexto linguístico de Guiné Bissau, o português tem sido considerado uma segunda língua, conforme Sani (1996), Fonseca (2012), o que nos leva a discutir, sob o enfoque da Teoria Gerativa, como ocorre a aquisição de uma L2. Sobre essa aquisição, algumas hipóteses têm sido levantadas, a saber:

- a) *Hipótese do acesso nulo*: os defensores dessa hipótese afirmam que os processos realizados para a aquisição da primeira e segunda língua são completamente diferentes, pois, enquanto a aquisição de L1 é guiada pela GU, a aquisição de L2 depende de outros fatores cognitivos e não da interferência da GU. Como representantes dessa hipótese, temos Bley-Vroman (1989) que defende a ideia de que os processos de aquisição de L1 e L2 são distintos, pois o Dispositivo de Aquisição da Linguagem não opera após o período da puberdade, assim, na aquisição da L2, não há restrição dos princípios da GU. Para esse autor, a aquisição da L2 pode não ser bem sucedida, pois não é uniforme em uma mesma comunidade linguística, já que há interferência de outros fatores. Ainda destacam-se Clahsen e Muysken (1986) e Clahsen (1988) ao afirmarem que a aquisição de L2 pode ser entendida como outro tipo de aprendizagem humana que não é guiada pelos princípios da GU; para eles, a aquisição da L2 dá-se por meio de outras

capacidades cognitivas, sendo assim, não há restrição da GU nem transferência da L1 no processo de aquisição da L2.

- b) *Hipótese do período crítico*: para essa hipótese, destaca-se Lenneberg (1967) ao defender que a aquisição da língua materna é inevitável por ser biologicamente prevista para acontecer, porém é cronologicamente limitada a um período crítico que vai até a puberdade. Conforme esse linguista, é a lateralização do cérebro que provoca a perda da capacidade de aquisição da linguagem na puberdade. Não obstante, Snow (1987) faz uma crítica a essa afirmação e admite que as evidências atestam a favor da presença de um fundamento biológico presente desde o nascimento – que é inato –, porém que não há uma indicação da idade ideal para a aquisição linguística completa. Além disso, pesquisadores atuais – em sua maioria – concordam que quanto mais aumenta a habilidade na língua materna, mais é difícil o processo de aprendizagem de uma segunda língua. Ainda sobre o período crítico, Kato (2003) afirma que existe uma idade crítica apenas no nível prosódico e fonológico da língua, mas que o nível sintático fica disponível no processo de aquisição de uma L2, independentemente da idade.
- c) *Hipótese do acesso total*: como defensores dessa hipótese, destacam-se Schwartz e Sprouse (1994, 1996) que asseguram que há transferência da gramática da língua materna, nos estágios iniciais, para a aquisição da segunda língua e posteriormente há acesso à GU; assim, o estágio inicial da L2 é o estágio final da L1. Ou seja, os valores paramétricos que estão determinados na L1 passam a constituir o estágio inicial da L2 que, a partir da exposição da criança ao *input*, desenvolve um novo sistema gramatical, reestruturando a gramática do estágio inicial. Nessa hipótese o desenvolvimento da gramática da L2 dá-se mediante ao uso da gramática da L2 seguido de falhas e que a partir dessas falhas, o aprendiz acessa os princípios da GU.
- d) *Hipótese do acesso parcial*: nessa hipótese, defende-se que há transferência parcial das propriedades da L1 para a L2 e, nos estágios subsequentes, a gramática é limitada pela GU. Vainikka (1989) e Strozzer (1992) sustentam que o estágio inicial da L2 é o estágio final da L1, porém apenas parte da gramática de L1 é considerada na formação do estágio inicial da L2; assim, a gramática inicial da L2

é incompleta no que se refere às categorias funcionais. Eubank (1993, 1994, 1996) propõe que o estágio inicial de L2 compreende toda a gramática da L1, exceto pelos valores dos núcleos funcionais, ou seja, todas as categorias que fazem parte da L1 são inicialmente transferidas para a gramática de L2, porém as particularidades das categorias funcionais são neutralizadas.

Os estudos de Flynn (1987; 1991), por exemplo, sustentam a hipótese de que os aprendizes de L2 são capazes de atribuir novos valores paramétricos no processo de construção da gramática de sua L2 quando os valores paramétricos da L1 não coincidem com as necessidades da segunda língua. Assim, assumimos para essa pesquisa que há acesso à GU e à gramática da língua materna no processo de aquisição de uma segunda língua, sendo somada a esse acesso a exposição ao *input*, com a finalidade de alcançar uma estrutura particular da gramática da segunda língua.

3.2.2.1 Aquisição da linguagem à luz da crioulistica

Lucchesi e Baxter (2009) definem a Transmissão Linguística Irregular como os processos históricos de contato maciço entre povos de línguas distintas, do século XVI ao XIX, por conta da ação do colonialismo europeu na África, América, Oceania e Ásia. Esses processos de colonização culminaram em um contato linguístico que gerou novas variedades linguísticas. Para formação dessas novas variedades, podemos dizer que havia a imposição da língua do grupo dominante, dos colonizadores, que era chamada de língua de superstrato ou língua-alvo, aos falantes das outras línguas. Em sua maioria, os falantes eram adultos, que eram obrigados a adquirir essa nova língua por questões sociais de subordinação.

Ao que sabemos, as variedades de português como L2 que se formaram nessas condições de contato viabilizaram o *input* para aquisição da língua materna para as novas gerações, na medida em que os grupos nativos iam supostamente abandonando suas línguas nativas. Conformes os autores, esse processo de nativização da língua do dominante ocorre de maneira irregular, pois os dados linguísticos primários a que as crianças são submetidas derivam, na realidade, de uma versão da segunda língua desenvolvida pelos falantes adultos das outras línguas nativas, que apresentam lacunas, defeitos e reanálises em relação aos

aspectos gramaticais. No processo de transmissão linguística geracional das línguas naturais, as crianças dispõem de dados primários mais completos, mesmo considerando a insuficiência do *input*, contrariando o que acontece no processo de transmissão linguística irregular. Assim, entendemos que o caminho para uma língua historicamente nova, uma língua crioula, é aberto pelo processo de transmissão linguística irregular ou que esse processo pode levar à formação de uma nova variedade histórica da língua de superstrato, que também apresentará processos de variação e mudanças motivados pelo contato entre as línguas. A variedade dessa língua de superstrato, na maioria das vezes, apresenta uma redução em sua estrutura gramatical, pois apenas os elementos essenciais às funções de comunicação básicas são mantidos, ou seja, a língua de superstrato transmite o léxico, enquanto a língua de substrato permanece com seus traços gramaticais. Assim, o resultado desse contato linguístico “massivo e abrupto”, conforme Lucchesi e Baxter (2009, p. 102), é um conjunto minimizado de itens lexicais que cada falante emprega de acordo com sua competência linguística nativa. É essa nova variedade que, com o passar do tempo, será passada para as demais gerações, que pode, como vimos, transformar-se em uma língua nova, uma língua crioula, ou em uma nova variedade da língua-alvo.

A partir do momento que uma língua crioula já está formada, ela desempenha as mesmas funções comunicativas e sociais das outras línguas naturais e é transmitida normalmente pelas as gerações. O que vai definir uma língua crioula é o processo acelerado de reestruturação da gramática na formação dessa língua, conforme os autores, “o grau de criouliização de uma língua pode ser mensurado relativamente com base no grau de reestruturação gramatical divergente dos modelos disponíveis na língua lexificadora”. (LUCCHESI e BAXTER, 2009, p. 119). Sendo assim, quanto mais as estruturas construídas no processo de reestruturação gramatical se distanciarem da língua-alvo, maior será o grau de pidginização e/ou criouliização.

Segundo Roberts (1997), com base na teoria de Princípios e Parâmetros, as línguas crioulas apresentam parâmetros de valores não marcados, mesmo quando na língua de superstrato esses parâmetros são marcados. À exemplo, temos a marcação positiva para o Parâmetro do Sujeito Nulo na língua portuguesa, enquanto que para o crioulo guineense, língua de base portuguesa, a marcação é negativa para esse mesmo Parâmetro. Sobre essa questão, também fundamentados em Roberts (1997), Lucchesi e Baxter (2009) afirmam:

Do ponto de vista da aquisição, “vê-se que só os valores marcados precisam ser expressos; no caso da não-expressão de um parâmetro, o *default* será o valor não marcado”. Desse modo, os crioulos apresentariam valores não marcados de parâmetros sintáticos em função das circunstâncias em que ocorre a sua gênese. Tendo como ponto de partida para a sua formação um *pidgin*, ou mesmo um mero leque de interlínguas (soluções individuais), o crioulo tende a se constituir com parâmetros fracos, em função da perda de morfologia que se dá na fase inicial do processo de transmissão linguística irregular. O estímulo-gatilho para a aquisição das gerações iniciais de falantes do crioulo pode ser “morfológica e sintaticamente defectivo em pontos cruciais” em relação a “certas propriedades requeridas para o desencadeamento de traços fortes”

Por fim, temos que essa afirmação corrobora com nossa hipótese de que os nossos dados do português guineense podem, por influência do crioulo guineense, ir de encontro ao que se verifica no português europeu, língua que, de acordo com Duarte (1995), como observamos no quadro 4, apresenta elevados índices percentuais em relação à distribuição do sujeito nulo. Ao que pensamos, os dados do PGB devem aproximar-se dos resultados do PB, encontrados por Duarte (1995), nos quais os sujeitos nulos predominam mediante algumas restrições.

3.2.2.2 Aquisição de português como L2 em Guiné Bissau

Como vimos, o português, embora seja a língua oficial do país, não é a língua materna em Guiné Bissau. O papel da língua materna é ocupado pelo crioulo guineense. É a essa língua que as crianças guineenses são expostas quando nascem, fixando o valor negativo para o PSN.

Couto e Embaló (2010) observam que o regime colonial e as circunstâncias de contato do português com as línguas africanas não facilitaram uma situação de plurilinguismo em Guiné Bissau. Os portugueses chegaram ao território que viria a ser a Guiné Bissau desde o século XVI, entretanto só começaram o processo de colonização a partir do século XIX. O contato dos portugueses com os nativos era restrito à área da escravização e não havia interesse, por parte dos portugueses, em ensinar a língua portuguesa aos nativos.

É importante salientarmos que o território ocupado pelos portugueses já era habitado por várias etnias e que essas etnias carregavam consigo culturas, línguas e costumes, havendo assim uma imensa diversidade de línguas africanas. Nesse contexto, não foram criadas situações que facilitassem o aprendizado do português pela maioria da população local, pois

não havia entre os colonizadores o desejo de integrar sua língua e cultura com a dos povos que já habitavam o lugar.

Com o decorrer do tempo, alguns nativos aprenderam o português, através do uso da língua e com a ajuda de alguns colonizadores, não havia inicialmente em Guiné Bissau (por parte do colonizador) o desejo de ensinar a língua, como houve no Brasil e os colonizadores dividiram-nos em assimilados e indígenas: estes não falavam o português e aqueles tinham se apropriado da língua e dos costumes portugueses, valendo frisarmos que o número de assimilados era muito restrito, pois poucos tiveram tal oportunidade de aprender a língua portuguesa, conforme Cruz (2013), Couto (1989)

Essa interação linguística entre colonizadores e colonos feita de maneira segregada, por parte dos colonizadores, trouxe para Guiné duas consequências: a) má distribuição do português por todo o território nacional e b) formação de um *pidgin* resultante do contato do com as línguas nativas, do qual nasceu o crioulo. Sobre essas consequências, Couto (1989, p. 108) afirma que:

Desde os primeiros anos da colonização, os colonizadores estiveram preocupados com a escravização dos africanos e não com lhes trazer a cultura europeia. Não houve a preocupação de ensinar o português aos nativos. Por causa disso e do fato de falantes de diversas línguas diferentes entre si terem que interagir neste espaço, ou seja, no "país" Guiné Bissau, foi surgindo naturalmente o crioulo como uma espécie de língua geral de comunicação interétnica.

De acordo com Duarte (2006), apenas no século XIX, quando se instalou o ensino do português em Guiné Bissau, o crioulo já era utilizado nos mais diversos domínios e encontrava-se estruturalmente bem mais estabilizado que a língua do colonizador, mesmo essa assumindo o papel de língua oficial do país após a independência.

Desde o tempo da colonização até os dias atuais, o ensino do português esteve restrito ao espaço escolar em Guiné Bissau. Conforme consta no quadro 1 do capítulo 1 desta dissertação, 80% da população guineense tem o crioulo guineense como língua materna, enquanto o português é falado apenas por cerca de 13% da população e é tido como segunda língua. Dentre as etnias que se destacam com maior número de falantes temos: balanta, fula, manjaco, mandinga, papel, mancanha, beafada, bijago, felupe, nalu, tanda, cocoli e susso. O

crioulo guineense, nesse contexto multilíngue, serve como um elo entre essas etnias. Sobre essa situação, destaca Sani (1996, p. 29):

Verifica-se que em Bissau, capital da Guiné, a maioria das crianças de cada grupo étnico fala o Crioulo como primeira língua e ao mesmo tempo fala a língua materna e aprende o Português na escola. Isso acontece principalmente quando há casamento entre pessoas de diferentes grupos étnicos e que passam a usar o Crioulo para falar com seus filhos, e não as línguas de suas etnias individuais. As crianças comumente convivem com a língua crioula mais do que com as línguas nativas.

Portanto, o português, embora seja língua oficial, ocupa o lugar de segunda língua em Guiné Bissau, sendo em geral aprendida na escola e falada por uma porcentagem muito baixa da população. Essa língua é usada por pessoas escolarizadas nos meios administrativos, políticos e nos principais meios de comunicação. Nessa acepção, constitui-se língua de prestígio, possibilitando o reconhecimento social associado à sua condição de língua da ascensão, do conhecimento e da cultura. Nesse país, foi promulgado um decreto (Decreto-lei – nº 7/2007 de 12 de novembro de 2007) que obriga à utilização do português em todas as instituições públicas; no ambiente escolar (dentro e fora da sala de aula), por exemplo, é obrigatório o uso do português. Apesar desse decreto, ainda se observa a grande taxa de analfabetismo 86%, conforme Couto e Embaló (2010).

Em linhas gerais, Guiné Bissau vive em um intenso contexto de multilinguismo: usam-se as línguas étnicas no contexto familiar quando pais e/ou avós ainda fazem uso dessas línguas (é um uso mais comum em áreas rurais); usa-se o crioulo guineense na comunicação social cotidiana (no ambiente familiar, nas feiras, nas ruas, ou seja, em contextos mais informais), e o português, nos contextos mais formais (escola, trabalho, repartições e instituições públicas). Sendo assim, apenas os guineenses escolarizados possuem a habilidade de usar a língua materna, o crioulo, e o português em diferentes locais e para fins distintos, ficando a grande maioria de falantes não escolarizados restritos ao uso de suas línguas étnicas e do crioulo guineense. Enfim, tomando por base que, em Guiné Bissau, o português é uma L2, acreditamos que a ocorrência de sujeitos nulos pode distanciar-se do PE.

4 METODOLOGIA

Este capítulo encarrega-se de descrever todos os passos utilizados na construção desta dissertação. Assim, detalharemos o perfil dos informantes que contribuíram fundamentalmente para este estudo, as fases da coleta de dados, a seleção e codificação das variáveis e, finalmente, as etapas que foram seguidas para a construção da dissertação e a realização da análise de dados que possibilitaram a conclusão deste trabalho.

4.1 MÉTODOS DE ABORDAGEM E PROCEDIMENTO

Para a realização desta pesquisa, utilizamos como método de abordagem o hipotético-dedutivo, visto que, conforme Lakatos e Marconi (2003), se analisa a percepção de uma lacuna nos conhecimentos, acerca da qual se formula hipóteses e, pelo processo de inferência dedutiva, testa-se a predição da ocorrência de fenômenos abrangidos pela hipótese. E como métodos de procedimento, o estatístico e o comparativo: este para estabelecer uma comparação entre as modalidades de fala e escrita e entre as variedades africanas do português com o PB no que se refere ao parâmetro do sujeito nulo, aquele para a realização da análise quantitativa do trabalho. Para tanto, analisamos dados da língua escrita extraídos de textos narrativos e dados da língua falada extraídos de entrevistas, que foram coletados de informantes guineenses e são-tomenses, alunos da UNILAB – CE.

4.2 PERFIL DOS INFORMANTES

Os nossos informantes são alunos de diversos cursos da UNILAB – CE, todos de Guiné Bissau e São Tomé. Após a aprovação do projeto desta pesquisa pelo Comitê de Ética da Universidade Federal de Pernambuco - UFPE, partimos para fazer a pesquisa de campo. Inicialmente, pensávamos em coletar dados de fala e de escrita de oitenta alunos africanos (40 de Guiné Bissau e 40 de São Tomé e Príncipe) de diferentes cursos (a saber: enfermagem, administração pública, sociologia, humanidades, agronomia, matemática, antropologia, engenharia de energias e ciências biológicas). Com a intenção de selecionarmos 10 homens e 10 mulheres dos primeiros períodos e 10 homens e 10 mulheres dos últimos períodos de cada

país, excetuando-se alunos do curso de Letras, pois entendemos que o grau de sensibilidade aos usos linguísticos é bem maior do que os demais alunos por estarem sendo formados para serem docentes de língua portuguesa.

Como é sabido que, em pesquisa de campo, é possível que algumas etapas programadas não ocorram como o previsto por razões de natureza diversa. Assim, ao irmos à UNILAB, foi informado que a instituição não recebeu alunos de São Tomé e Príncipe em 2018, conseqüentemente, não foi possível encontrarmos alunos são-tomenses ingressantes nesse ano. Nesse caso, assumimos os primeiros períodos como: primeiro, segundo, terceiro e quarto semestres e os últimos períodos a partir do sexto semestre. Também não obtivemos o número esperado de mulheres guineenses, o que nos levou a reduzir nosso total de 20 informantes (homens e mulheres de cada país) para 12 informantes (homens e mulheres de cada país), com o propósito de equiparmos nossos resultados. Abaixo fizemos um quadro elencando a quantidade de informantes por cada curso selecionado, diferenciando o sexo e a variedade africana do português, com o propósito de termos um mapeamento mais detalhado dos sujeitos da pesquisa, vejamos:

Quadro 16 - Relação dos informantes da pesquisa de acordo com curso, sexo e variedade africana do português.

Cursos	Português São-tomense		Português Guineense	
	Masculino	Feminino	Masculino	Feminino
Engenharia de Energias	6	2		
Antropologia			1	
Matemática			2	
Agronomia	3	3	4	
Humanidades			2	2
Sociologia		1	2	2
Administração Pública	2	3	1	2
Enfermagem	1	2		6
Ciências Biológicas		1		

Portanto, nossa pesquisa abrange 12 homens e 12 mulheres guineenses e 12 homens e 12 mulheres são-tomenses dos primeiros e últimos períodos dos diversos cursos superiores oferecidos na UNILAB – CE, totalizando 48 informantes. Vale ressaltarmos que todos esses informantes moravam em zonas urbanas em seus países de origem e ingressaram nessa universidade após sua aprovação em uma das etapas do processo seletivo, a saber: uma redação feita em seu país de origem.

Todos os colaboradores desta pesquisa leram e assinaram o Termo de Consentimento Livre e Esclarecido – TCLE, que passou pelo Comitê de Ética da UFPE. Nesse documento, consta que todas as informações desta pesquisa são confidenciais e divulgadas apenas em eventos ou publicações científicas, não havendo identificação dos voluntários, a não ser entre os responsáveis pelo estudo, sendo assegurado o sigilo sobre a sua participação e ficando os dados coletados nesta pesquisa (gravações de entrevistas e textos narrativos) armazenados em nosso computador pessoal.

4.3 COLETA DE DADOS

Nossa coleta de dados aconteceu na cidade de Redenção-CE, localizada a um pouco mais de 50km da capital, Fortaleza. Esse município recebeu o título de primeira cidade do Brasil a abolir a escravidão. Existem dois museus na cidade que mostram as memórias daquele tempo. A UNILAB, universidade onde coletamos os dados é uma instituição federal de integração com outros países, principalmente, com os países africanos onde se fala português. Criada em 2010, tem como objetivo principal formar recursos humanos para contribuir com os países lusófonos. Conforme o perfil da universidade UNILAB (2018), temos:

De acordo com a legislação, a Unilab tem como objetivo ministrar ensino superior, desenvolver pesquisas nas diversas áreas de conhecimento e promover a extensão universitária, tendo como missão institucional específica formar recursos humanos para contribuir com a integração entre o Brasil e os demais países membros da Comunidade dos Países de Língua Portuguesa (CPLP), especialmente os países africanos, bem como promover o desenvolvimento regional e o intercâmbio cultural, científico e educacional.

Sendo assim, recebe tanto estudantes brasileiros quanto estudantes estrangeiros, principalmente africanos, o que justifica a nossa escolha para a coleta de dados.

Resolvemos trabalhar com as modalidades de escrita e de fala, para assim, fazermos um comparativo entre essas modalidades no que se refere ao uso do PSN. Trabalhamos com os mesmos informantes para obtermos os dados tanto de escrita quanto de fala. É importante frisar que os informantes selecionados estão entre os primeiros e últimos períodos de seus cursos, pois existe a necessidade de se fazer uma análise comparativa entre os dados para constatar se os informantes dos primeiros períodos tendem a produzir mais sujeitos nulos em decorrência de ter o PE como exemplo da norma culta do português, enquanto que os informantes dos últimos períodos tendem a usar mais sujeitos plenos em decorrência do frequente contato com o PB. Para tanto, fomos duas vezes à UNILAB – CE: na primeira vez, fizemos contato, conhecemos nossos informantes (não aplicamos uma ficha que registrasse o perfil social de nossos informantes, contudo obtivemos as informações da localidade em que eles moravam, se áreas urbanas ou rurais, e se estudaram em escolas públicas ou privadas) e fizemos a coleta dos dados escritos e, no segundo momento, retornamos à universidade, contactamos os mesmos informantes e coletamos os dados de fala.

Os dados de escrita foram obtidos a partir de textos narrativos elaborados pelos alunos. Para garantir uniformidade, a estória desenvolvida pelos informantes teve um início elaborado pelo pesquisador para que eles dessem continuidade sem estipularmos o número de linhas (anexaremos um modelo do nosso material para coleta de dados ao fim do nosso trabalho). Acrescentamos que a escolha de um gênero textual predominantemente narrativo se deu, pois, queríamos instigar a escrita livre dos alunos, menos policiada, para que eles se envolvessem com a narrativa. Todos esses textos encontram-se digitalizado e, dentre eles, foram selecionados aqueles textos que apresentaram um maior número de linhas. Após selecionados esses textos, voltamos à Redenção e entramos em contato com os informantes para realizarmos as entrevistas informais.

Para os dados de fala, realizamos entrevistas informais com cada informante individualmente, tendo duração média de 7 a 12¹⁴ minutos (embora o tempo não seja tão longo, o fenômeno em estudo é recorrente). Essas entrevistas foram feitas com base em um roteiro de perguntas (que poderá ser encontrado em nossos anexos), realizadas tanto em

¹⁴ Em geral, as entrevistas têm um tempo maior, porém, o fenômeno analisado é recorrente, assim, não foi necessário fazer longas entrevistas.

espaço aberto (pátio da universidade, restaurante universitário) quanto em espaços fechados (biblioteca, salas de aula, casa dos informantes), estando armazenadas em gravador MP4. Não tínhamos como estipular um único local para as entrevistas, pois os alunos têm uma rotina preenchida de atividades acadêmicas e muitos dos cursos acontecem em período integral na universidade. Assim, onde o aluno estivesse, íamos e fazíamos a entrevista. Concluída a coleta, passamos para a transcrição dos dados de fala, usando como ponto de referência o sistema ortográfico.

4.4 CONSTITUIÇÃO DO *CORPORA*

Para a constituição dos *corpora* da língua falada e escrita de alunos dos primeiros e últimos períodos de diversos cursos da UNILAB, elegemos contextos declarativos finitos contendo sujeitos nulos e plenos. Observem-se alguns exemplos do PGB (cf. (26)) e do PST (cf. (27)) extraídos dos *corpora*:

(26)a. “**pro** Tive o privilégio de ser aprovada para a bolsa da UNILAB”. (I.3, GB, F, PI)

a’. “**Eu** adoro meu país”. (I.1, GB, F, PI)

b. “**pro** Aprendi a língua portuguesa ao longo do tempo” (I.6, GB, M, PF)

b’. “A língua que **eu** mais tenho domínio na minha vida é a língua crioula”. (I.6, GB, M, PF)

(27)a. “**pro** Temos grandes pontos turísticos que encantam” (I.3, STP, M, PI)

a’. “Lá **eu** era graduando em relações internacionais” (I.2, STP, M, PI)

b. “**pro** Começo já dizendo que a língua portuguesa é minha língua materna” (I.1, STP, F, PI)

b’. “**Eu** tenho dificuldade de falar língua crioula” (I.6, STP, F, PF)

Obtivemos para o *corpus* do PGB um total geral de 1.220 dados e, para o *corpus* do PST, 1.043 dados. Observe-se o quadro, a seguir, com o quantitativo de dados já obtidos, tomando por base a modalidade de língua em cada país analisado:

Quadro 17 - Quantitativo de dados obtidos de acordo com a modalidade de língua e a variedade africana do português.

Modalidade de língua	Português de Guiné Bissau	Português de São Tomé e Príncipe
Língua Escrita	578	458
Língua Falada	642	585

4.5 SELEÇÃO E CODIFICAÇÃO DAS VARIÁVEIS

Elaborados os *corpora*, elencamos as variáveis linguísticas e extralinguísticas para a testagem das hipóteses de trabalho, visando à codificação de todos os dados selecionados. Para a seleção das variáveis linguísticas, apoiamo-nos nas variáveis estudadas por Duarte (1995). Ela trabalhou com as:

- a) Pessoa gramatical e desinência verbal
- b) A estrutura da oração e seu estatuto sintático dentro do período
- c) O traço do referente de terceira pessoa
- d) O duplo sujeito

É importante salientarmos que, na escolha das variáveis extralinguísticas, não trabalhamos a área de moradia dos informantes em seu país de origem nem a língua que eles consideravam materna. Isso aconteceu porque a) todos os estudantes moravam em áreas urbanas em Guiné Bissau ou São Tomé e Príncipe e b) todos os falantes guineenses declararam que têm como língua materna o crioulo e todos os são-tomenses, o português. Nesse sentido, elencamos como variáveis extralinguísticas: a) o país de origem, b) o período do curso e c) o sexo.

Apresentamos, a seguir, a variável dependente e as variáveis independentes (linguísticas e extralinguísticas) com seus respectivos fatores antecedidos de um código (que se encontra em negrito) a partir do qual os dados foram codificados para posterior tratamento quantitativo:

VARIÁVEL DEPENDENTE

0- Sujeito nulo:

(28) “**pro** Tenho 24 anos de idade” (I.2, GB, F, PI)¹⁵

1- Sujeito pleno:

(29) “**Eu** tinha uma vida tranquila” (I.4, GB, F, PI)

VARIÁVEIS EXTRALINGUÍSTICAS

1) País de origem:

G- Guiné Bissau

S- São Tomé e Príncipe

2) Período do curso:

P- Primeiros períodos

U- Últimos períodos

¹⁵ I = Informante; GB = Guiné Bissau; STP = São Tomé e Príncipe; F = Feminino; M = Masculino; PI = Períodos iniciais; PF = Períodos finais.

3) Sexo:

H- Homem

M- Mulher

VARIÁVEIS LINGUÍSTICAS

– A pessoa gramatical e a desinência verbal

3- Primeira pessoa do singular:

Ex.: (30) “**Eu** morava com meus pais” (I.2, GB, F, PF)

5- Segunda pessoa do singular:

Ex.: (31) “**pro** Entendes o que eu falo?” (I.3, STP, F, PI)

6- Terceira pessoa do singular:

Ex.: (32) “**Ela** possui floresta virgem” (I.5, STP, M, PF)

7- Primeira pessoa do plural:

Ex.: (33) “**pro** Vivíamos no mesmo bairro” (I.2, STP, F, PI)

8- Segunda pessoa do plural

– Não houve ocorrências

9- Terceira pessoa do plural

Ex.: (34) “**Eles** são tudo pra mim” (I.1, GB, M, PF)

– A estrutura da oração e seu estatuto sintático dentro do período

E- Oração Encaixada:

Ex.: (35) “Minha mãe aceitou o que meu pai disse que era pra fazer, deixou eu vim para o Brasil”. (I.5, GB, M, PF)

O- Outras Orações:

Ex.: (36) “Eu amo meu país, amo minha família, meus amigos e todo povo guineense” (I.2, GB, F, PI)

– O traço do referente de terceira pessoa

A- [+] animado:

Ex.: (37) (referente: minha mãe) “foi quem me incentivou a fazer a inscrição da UNILAB, ela dizia que eu não podia perder essa chance, fez de tudo para que eu pudesse vim”. (I.3, STP, M, PI)

Z- [+] animado e [+] genérico:

Ex.: (38) (referente: meus amigos) “que já estavam aqui no Brasil, que me falaram sobre a oportunidade de vir estudar aqui, eles me ajudaram muito” (I.2, GB, M, PI)

I- [-] animado:

Ex.: (39) (referente: o meu país) “é bom demais de se viver, não tem violência, é considerado sem racismo, pra mim é o melhor lugar” (I.4, STP, F, PF)

– Duplicação do Sujeito Pleno¹⁶

2- Sem duplicação de sujeito pleno:

Ex.: (40) “Principalmente a minha mãe que fez tudo que podia para que eu pudesse estudar”. (I.3, STP, F, PI)

4- Com duplicação do sujeito pleno:

Ex.: (41) “A minha terra, ela é maravilhosa”. (I.1, STP, M, PI)

Dessa forma, selecionamos nossas variáveis e passamos para a codificação dos dados. Toda codificação foi feita de maneira mnemônica para facilitar nossas análises no decorrer da pesquisa.

4.6 TRATAMENTO QUANTITATIVO DOS DADOS

Realizada a codificação, os dados foram submetidos a tratamento quantitativo através de sua rodada no programa computacional GoldVarb X que apresentou os resultados estatísticos e probabilísticos, levando em conta o nível de significância das variáveis. Obtidos esses resultados, seguimos para a análise linguístico-comparativa. É importante salientar que primeiramente rodamos os dados do PGB e do PST separados, com o intuito de obter percentuais específicos dos países para ter a liberdade de confrontar esses dados com pesquisas já existentes com o PE e o PB.

¹⁶ Não obtivemos nenhuma ocorrência dessa estrutura nos dados coletados de São Tomé e Príncipe, tanto nos dados de escrita quanto nos de fala.

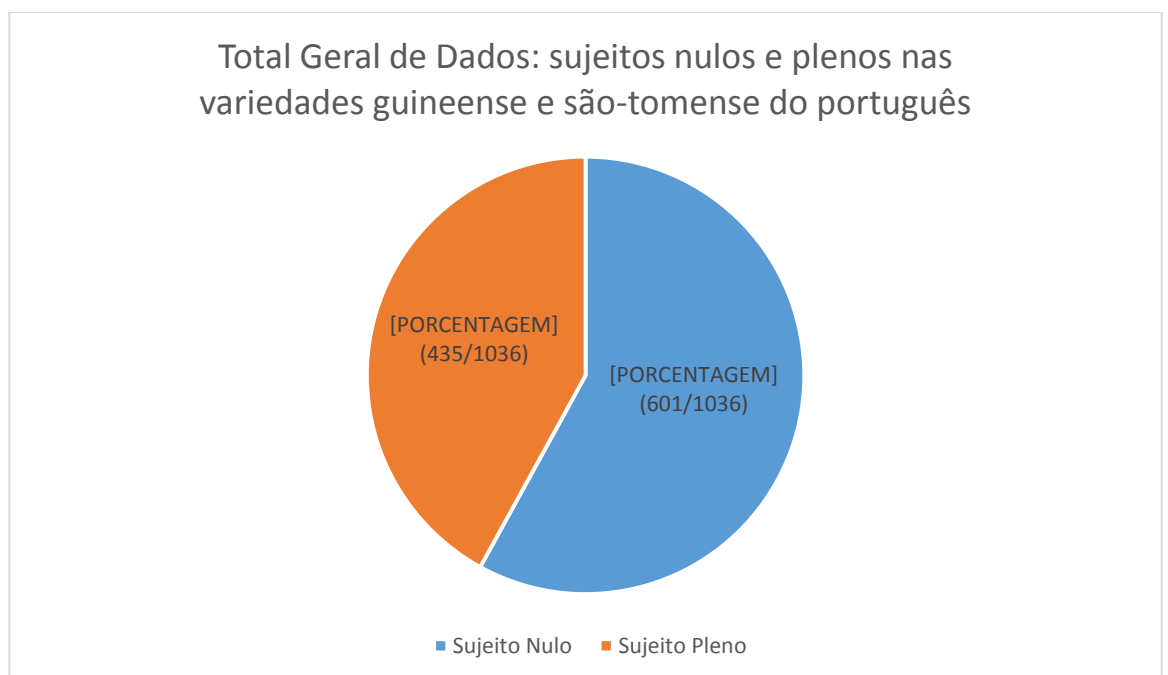
5 ANÁLISE DOS DADOS DE ESCRITA

Este capítulo dará conta da análise de dados de escrita. Na primeira subseção, faremos a comparação dos resultados quantitativos das duas variedades (PGB e PST) obtidos nas rodadas do programa Goldvarb X e sintetizaremos as semelhanças e diferenças das variedades. Posteriormente, em uma segunda subseção, faremos a análise linguística dos dados escritos do PGB e, na terceira subseção, dos dados do PST.

5.1 COMPARANDO GUINÉ BISSAU E SÃO TOMÉ E PRÍNCIPE: QUANTITATIVO GERAL DOS DADOS

Temos, em nossos dados de escrita, um quantitativo geral de 1036 sentenças, sendo 578 de Guiné Bissau e 458 de São Tomé e Príncipe. \213Em conformidade com esse quantitativo e considerando a variável dependente (sujeito nulo e sujeito pleno), o percentual revela uma maior ocorrência de sujeitos nulos (58%), ao contrário dos plenos (42%), obtendo uma diferença sutil de 16%. Vejamos o gráfico a seguir:

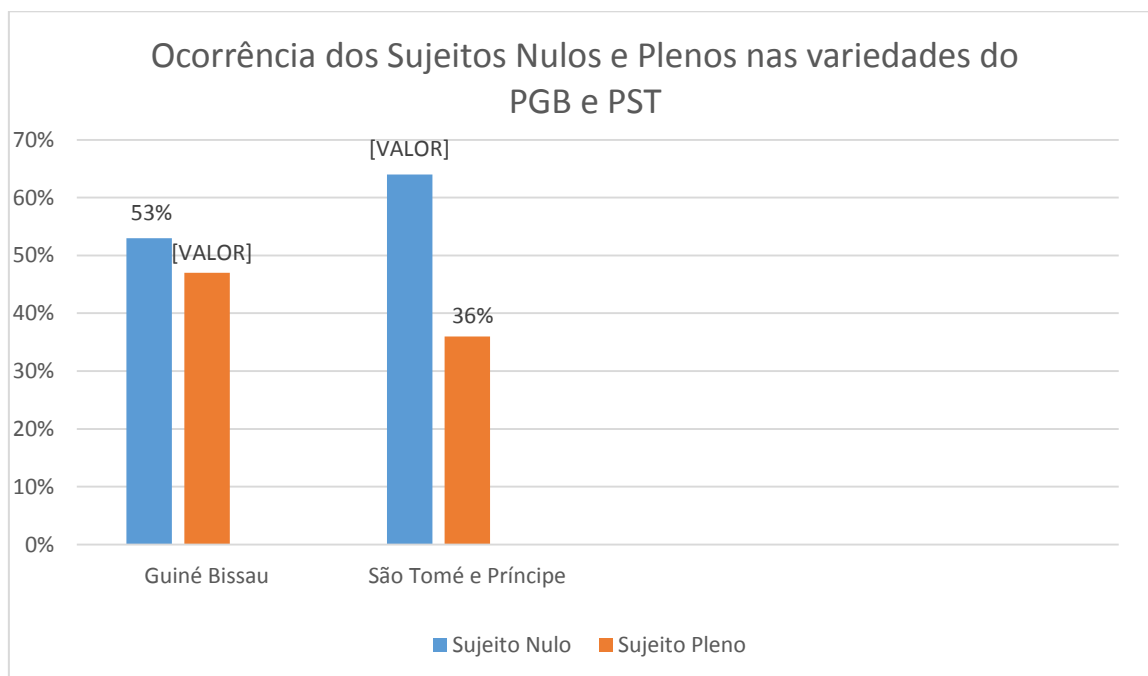
Gráfico 3 - Distribuição geral e percentual de sujeitos nulos e plenos nas variedades guineense e são-tomense do português.



Ao que tudo indica, esse resultado mostra-nos, a princípio, que o PGB e o PST encontram-se no ‘meio termo’ entre o PB e o PE, pois apresenta um percentual de sujeitos nulos maior que o encontrado para o PB e menor que o encontrado para o PE. Obtivemos uma diferença sutil de apenas 16% ao analisarmos, de maneira geral, a distribuição de sujeitos nulos e plenos nos dados das duas variedades. Em relação à variedade brasileira do português, Duarte (1995), como pudemos observar no gráfico 1, mostra-nos que em 1992 havia apenas 26% das ocorrências de sujeito nulo em PB.

Esses resultados são menores que os encontrados para o PE, pois, no que concerne à frequência de uso do sujeito nulo, como vimos no quadro 4, quando tratamos do PE, Duarte (1995) mostra que todas as pessoas gramaticais demonstram preferência pelo uso do sujeito nulo, tendo um total de 66% de nulos em concorrência com 34% de plenos. Diante desses resultados, ratificamos que, ao que nos parece, o PST e o PGB estão a meio caminho entre o PB e o PE. Vejamos a seguir como os sujeitos nulos e plenos distribuíram-se, tomando por base o país de origem dos alunos:

Gráfico 4 - Ocorrência de sujeitos nulos e plenos nas variedades do português de Guiné Bissau e São Tomé e Príncipe (%).

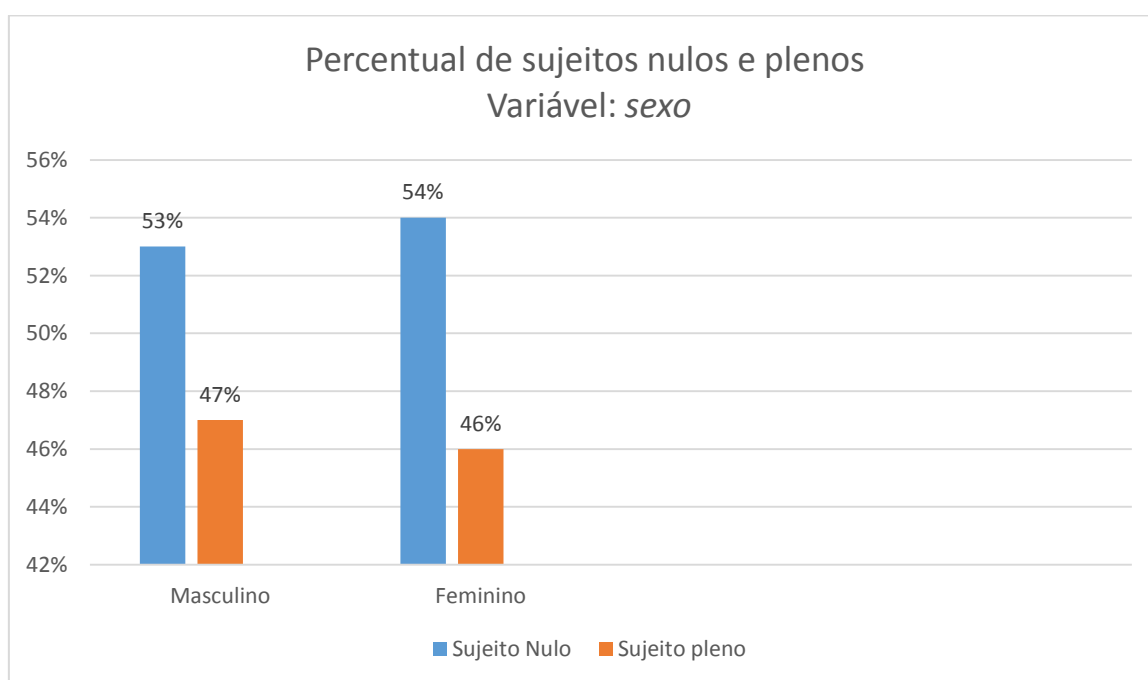


Conforme o gráfico 4, percebemos que a ocorrência de sujeitos nulos é maior em ambas variedades, sendo de 53% (308) no PGB e 64% (293) no PST, enquanto o sujeito pleno tem ocorrência de 47% (270) em Guiné Bissau e 36% (165) em São Tomé e Príncipe. Esse

resultado parece apontar para o que afirmamos em relação à aquisição da linguagem: a diferença de 11% da ocorrência de sujeitos nulos entre uma variedade e outra pode ser atribuída ao fato de o português ser adquirido como segunda língua em Guiné Bissau e, em sua maioria, como língua materna em São Tomé e Príncipe, acrescentando a esse fato o maior contato linguístico do português com o crioulo guineense em GB, pois sabemos que mais de 80% da população têm essa língua não-pro-drop como materna.

No que se refere à variável *sexo*, elaboramos gráficos separados para mostrarmos os resultados das duas variedades. Os resultados de Guiné Bissau mostram que 48% das mulheres realizaram o sujeito nulo, já os homens, 52%, sendo, portanto, mínima a diferença percentual entre eles (4%):

Gráfico 5 - Percentual de sujeitos nulos e plenos, realizados no PGB, tomando por base a variável *sexo*.



Como bem sabemos, as mulheres têm um papel significativo para os processos variáveis de diferentes níveis linguísticos, por isso, por vezes, lideram um processo de mudança linguística. Sabemos também, de acordo com Paiva (2017), que as mulheres possuem um papel mais conservador em relação às formas menos prestigiadas, assim, preferindo as formas mais prestigiadas da língua. Porém, analisando em nossa pesquisa o uso de sujeitos nulos e plenos, não podemos afirmar que se trata da existência de uma variante de

maior ou menor prestígio social, sendo assim, tanto homens quanto mulheres apresentam um comportamento linguístico muito semelhante.

Entendemos também que a diferença mínima de 1%, no que concerne à ocorrência de sujeitos nulos, entre o sexo masculino e o sexo feminino, pode ser atribuída ao fato de que, embora o crioulo guineense seja mais falado que o português, os falantes usam mais nulos em decorrência possivelmente do contexto que exige maior grau de monitoramento do uso linguístico na escrita por se tratar de uma situação de exame. Embora tivéssemos explicado claramente que eles não seriam avaliados pelo texto que estavam fazendo, mas entendemos que a escrita implica em um contexto de maior monitoramento linguístico.

Identificamos abaixo alguns exemplos extraídos do *corpus* do PGB que evidenciam o uso de sujeitos plenos e nulos por homens e mulheres:

- *Sexo masculino:*

(42) a. “**Eu** amo meu país”. (I.3, GB, M, PI)

b. **pro** “Sou de uma família numerosa”. (I.1, GB, M, PI)

- *Sexo feminino:*

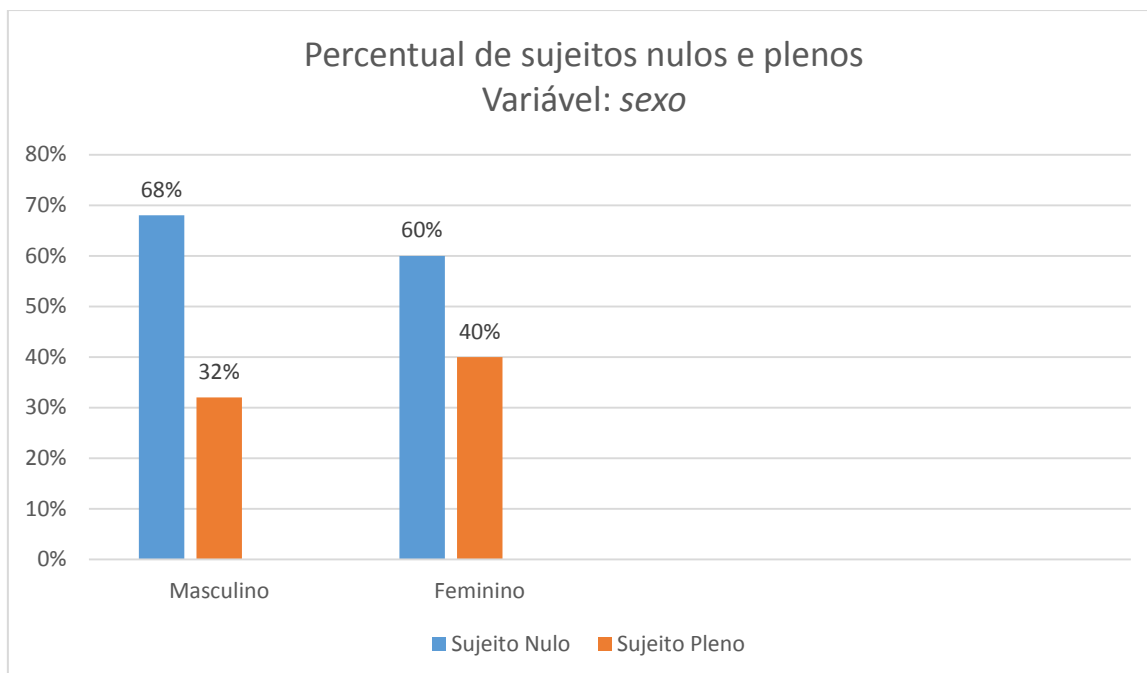
(43) a. “**Eu** nasci na capital Bissau”. (I.2, GB, F, PI)

b. **pro** “Estudei numa escola privada durante o ensino médio”. (I.1, GB, F, PI)

Comparando nossos resultados referentes à variável *sexo* aos de outros estudos, temos Duarte (1995) que, em relação à ocorrência de sujeitos nulos, depara-se com 34% dessas ocorrências para homens e 25% para mulheres, assim, apresentando uma diferença de 9% entre homens e mulheres. Nossos resultados também confirmam que não há grandes diferenças quanto ao uso de sujeitos nulos por homens e mulheres, pois não há para esse uso um estigma social.

Em relação aos resultados de São Tomé e Príncipe, no que concerne à variável *sexo*, temos também uma diferença sutil entre as ocorrências de sujeitos nulos e plenos. Vejamos:

Gráfico 6 - Percentual de sujeitos nulos e plenos, realizados na variedade são-tomense do português, tomando por base a variável sexo.



Atribuímos ao resultado acima o fato de que, em sendo o português não mais uma L2 para a maioria dos falantes são-tomenses, de acordo com Hagemer (2009), homens e mulheres são mais conservadores ao fazerem uso de estruturas mais próximas à norma europeia do português, confirmando-se assim nossa predição inicial de que os falantes são-tomenses fariam uso semelhante aos falantes do português europeu, no que concerne às ocorrências do sujeito nulo.

Observamos abaixo alguns exemplos extraídos do *corpus* da pesquisa que evidenciam o uso de sujeitos plenos e nulos por homens e mulheres são-tomenses:

- *Sexo Masculino:*

(44) a. “**Eu** tinha uma vida maravilhosa em São Tomé”. (I.1, STP, M, PI)

b. *pro* “Temos grandes pontos turísticos que encantam”. (I.3, STP, M, PI)

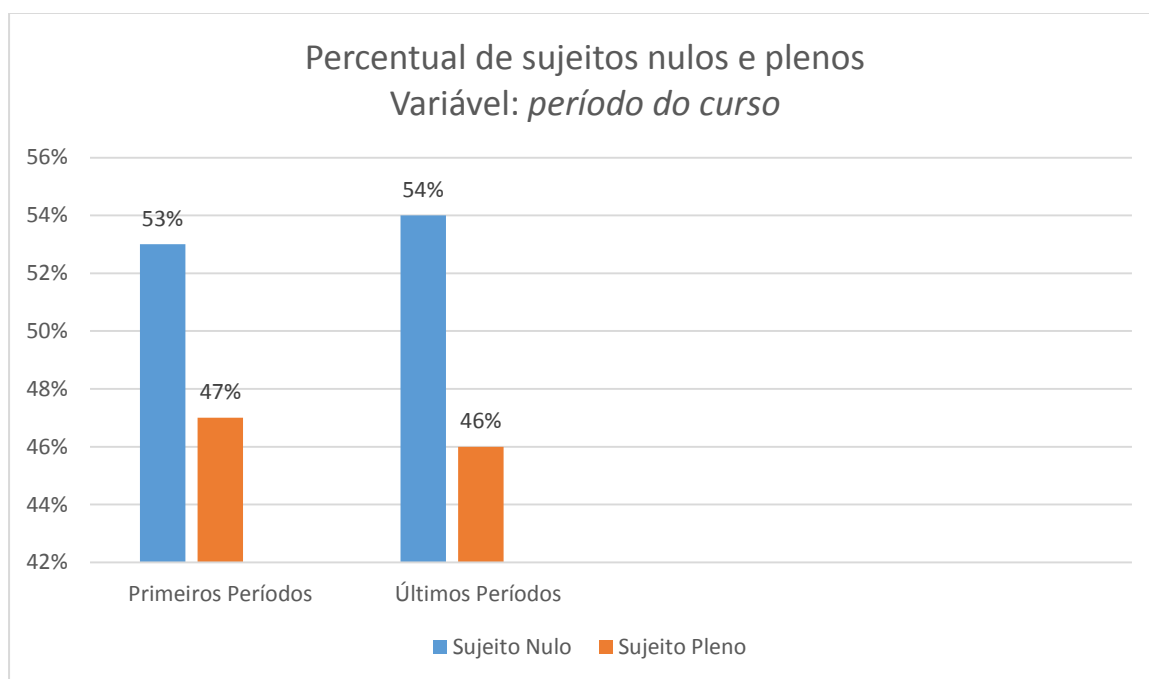
- *Sexo Feminino:*

(45)a. “**Eu** sempre soube que um dia iria sair do país”. (I.1, STP, F, PI)

b. **pro** “Tive o privilégio de nascer em São Tomé e Príncipe”. (I.1, STP, F, PI)

Para nossa pesquisa, o tempo de estadia dos informantes no Brasil é fundamental para que façamos a análise da interferência do português brasileiro nas variedades em questão. Como visto anteriormente, tanto em São Tomé e Príncipe quanto em Guiné Bissau a língua-alvo é o português europeu, e, como sabemos, é uma língua de sujeito nulo consistente. Por sua vez, o PB, língua em que os estudantes são-tomenses e guineenses entram em contato a partir do momento que chegam no Brasil, é uma língua de sujeito nulo parcial, ou seja, admite sujeitos nulos em contextos restritivos, conforme observam Cyrino, Duarte e Kato (2000) e Kato e Duarte (2014). Assim, aventamos a hipótese de que os informantes que estão no Brasil a mais tempo e consequentemente tem maior tempo de contato com PB possam apresentar em sua variedade africana do português interferências do PB no que concerne ao uso do sujeito nulo. Dentre essas interferências, esperamos encontrar o uso de expressões pronominais como *você* e *a gente*, a duplicação do sujeito e possível enfraquecimento da morfologia de flexão verbal. Para tanto, mostramos os resultados obtidos para Guiné Bissau e São Tomé em relação ao período de tempo que os estudantes estão na universidade:

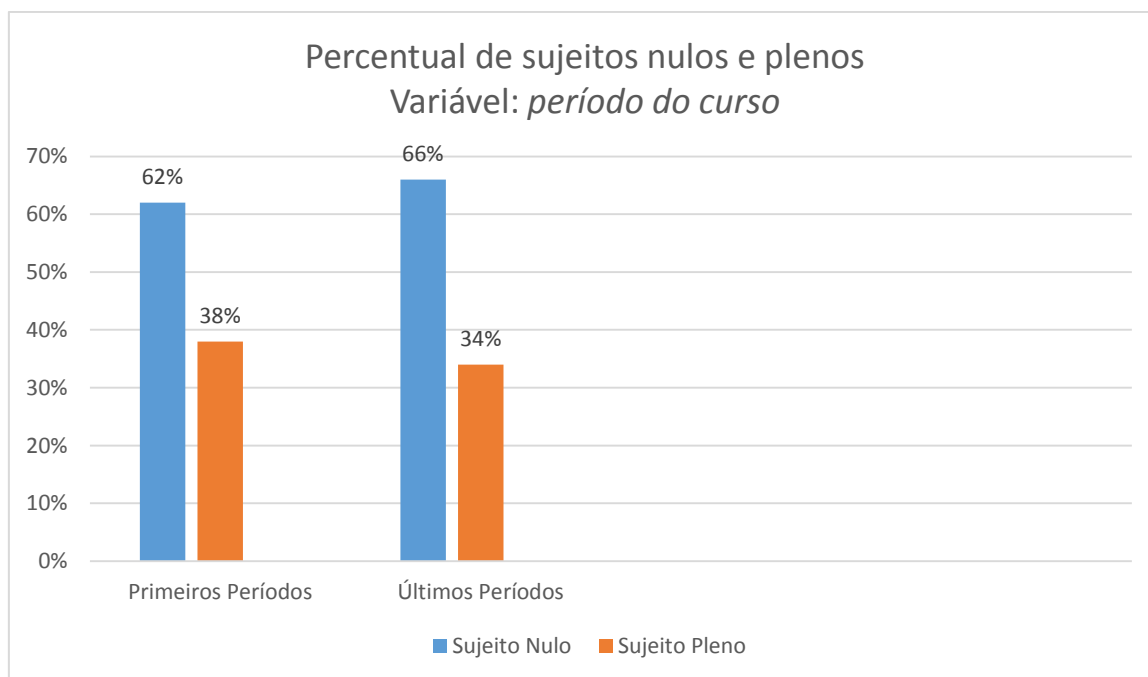
Gráfico 7 - Percentual de sujeitos nulos e plenos, realizados na variedade guineense do português, tomando por base a variável período do curso.



No gráfico 7 acima, podemos observar uma diferença mínima, em relação ao uso do sujeito foneticamente expresso entre os informantes guineenses recém-chegados ao Brasil (primeiros períodos) e os informantes que estão há mais de três anos país (últimos períodos). Além disso, ainda temos um decréscimo de 1% das ocorrências de sujeitos plenos: de 47% desses sujeitos nos primeiros períodos para 46% nos últimos períodos. Esse resultado vai de encontro à nossa hipótese de uma possível interferência do PB sobre o PGB, pois não houve um crescimento considerável do número de ocorrências de sujeitos foneticamente expressos por alunos que estão há mais tempo convivendo com o PB, língua cujos sujeitos nulos encontram-se submetidos a restrições.

Em relação à São Tomé e Príncipe, os resultados não foram muito diferentes: houve um aumento percentual no uso de nulos (4%) quando comparados os primeiros e últimos períodos na universidade. Vejamos:

Gráfico 8 - Percentual de sujeitos nulos e plenos, realizados na variedade são-tomense do português, tomando por base a variável período do curso.



No gráfico 8, não muito diferente do que aconteceu com o PGB, temos um aumento de 4% nas ocorrências de sujeitos nulos, registrando 62% nos primeiros períodos e 66% nos últimos. Novamente, não é possível percebermos uma interferência do PB no português escrito por estudantes africanos da UNILAB, visto que os resultados referentes ao uso do sujeito nulo não sofreram grandes mudanças, inclusive, tiveram pequenos aumentos

percentuais. Todos os falantes entrevistados têm o PE como língua materna e que, por ser também oficial, é considerada de prestígio por promover a ascensão social de seus falantes em detrimento do santomé. Conforme Petter (2015, p. 218), essa “atitude positiva em relação à língua oficial leva à rejeição por grande parte dos pais, ao ensino em línguas africanas, pois eles desejam que seus filhos sejam proficientes nas línguas que promovem a ascensão social”. Em consequência de atitudes como essa, temos o fato de o PST não ser mais considerada como segunda língua por nossos informantes. Em nossos dados da modalidade escrita da língua, alguns informantes relataram casos de proibição por parte de seus pais ao usar o crioulo falado por seus avós, o santomé, fato este que corrobora o que foi dito por Petter (2015). Vejamos o depoimento de um dos informantes:

- (46) “Em casa e na escola devemos sempre falar português, moramos eu, meu pai, minha mãe e minha avó, ela fala crioulo, mas não sei falar, quando era miúda que tentei aprender, minha mãe castigou-me”. (I3, STP, F, PI)

No que concerne às diferenças percentuais encontradas nas variedades guineenses e são-tomenses, entre os primeiros e últimos períodos, vê-se um percentual maior para informantes recém-chegados de São Tomé e Príncipe (62%) em relação ao percentual encontrado para informantes recém-chegados de Guiné Bissau (53%). Atribuímos esses resultados ao tipo de aquisição do português que se dá de maneira distinta nos dois países.

Em linhas gerais, os resultados quantitativos apresentados nesta subseção revelam que o PST é mais próximo à norma do PE por legitimar mais sujeitos nulos (64%) do que o PGB (53%) que se afasta um pouco desta.

5.2 AS VARIÁVEIS LINGUÍSTICAS: GUINÉ BISSAU

5.2.1 Pessoa gramatical e desinência verbal

Observemos no quadro abaixo, a frequência de sujeitos nulos encontrados em relação à referência semântica pronominal, pessoa do discurso¹⁷:

¹⁷ Ao cruzarmos o tempo de estadia dos informantes no Brasil com a pessoa do discurso, não obtivemos diferenças significantes entre os resultados, crescendo apenas 1% de diferença entre os primeiros e últimos

Quadro 18 - Ocorrência de sujeito nulo, na modalidade escrita da língua, na variedade guineense, segundo a pessoa do discurso

Pessoas do Discurso	Sujeito Nulo	Sujeito Pleno	Peso Relativo
1ª pessoa do singular	76% (203)	24% (64)	0, 61
3ª pessoa do singular	68% (169)	32% (78)	0,54
1ª pessoa do plural	73% (11)	27% (4)	0,59
3ª pessoa do plural	68% (32)	32% (15)	0,53

Analisando o quadro 18, chama-nos atenção a grande ocorrência da primeira e da terceira pessoas do singular, somando mais da metade dos dados analisados. Entendemos que esse número exacerbado de ocorrências de sujeitos nulos e plenos especificamente nessas duas pessoas do discurso aconteça devido à proposta de narrativa que foi escolhida para coletar os dados, visto que em uma narrativa, geralmente, predominam a primeira e/ou terceira pessoa do discurso, a depender do foco narrativo. Como nossa proposta de narrativa dava margem ao discurso direto, obtivemos esse posicionamento por parte dos informantes. Dito isso, seguimos para nossa análise.

Como podemos observar, há predominância de sujeitos nulos tanto na primeira pessoa do singular (76%) quanto na primeira pessoa do plural (73%). Esse resultado aproxima o PGB, em sua modalidade escrita da língua, do PE. Conforme Duarte (1995), o PE tem preferência pelo sujeito nulo em todas as pessoas do discurso, apresentando percentual de 60% de sujeitos nulos nas primeiras pessoas do singular e do plural.

Pela escala de referencialidade proposta por Kato, Duarte e Cyrino (2000), as pessoas que menos apresentam sujeito nulo no PB são as 1ª e 2ª pessoas do discurso, como vimos, ao analisarmos os dados da modalidade escrita da língua, o PGB apresenta um percentual alto de sujeitos nulos nas primeiras pessoas, assim, afastando-se do PB e contrariando nossa predição inicial de uma interferência do PB sobre o PGB. É necessário pontuarmos que essa constatação, até o presente momento, restringe-se aos dados da língua escrita, pois,

períodos para a primeira pessoa do singular, em relação ao sujeito nulo. As demais pessoas do discurso apresentaram resultados menores que 1%. Sendo assim, optamos por apresentar e analisar as variáveis separadamente.

ressaltamos sobre a diferença quanto ao uso de nulos entre a língua escrita e a língua falada, sendo aquela mais favorecedora de nulos por ser mais conservadora

Abaixo, nas sentenças (47) e (48), podemos observar alguns exemplos dos dados de 1ª pessoa do singular e do plural, respectivamente:

(47) a. “**Eu** nasci na capital Bissau” (I1, GB, F, PI)

b. “**pro** Morava com meu pai, minha madrastra e minhas duas irmãs” (I1, GB, F, PI)

(48) a. “**Nós** como alunos sentimos essa grande dificuldade de ensino público” (I5, GB, M, PF)

b. “**pro** Comemos em cada casa em que entramos, **pro** pegavamos frutas onde quer que **pro** entramos **pro** tínhamos mães e pais por todo o arredores” (I6, GB, M, PI)

Ainda sobre a primeira pessoa do plural, destacamos que, dentre todos os dados analisados da modalidade escrita da língua, apenas uma informante, dos últimos períodos, fez o uso da expressão pronominal *a gente* que foi usada duas vezes no mesmo texto. Vejamos as ocorrências em (49):

(49)a. “No começo não foi fácil mas **a gente** vive um dia de cada vez”. (I4, GB, F, PF)

b. “De como **a gente** vive, nossas culturas, condições económicas...” (I4, GB, F, PF)

Dessa forma, compreendemos que o não uso da expressão pronominal *a gente* por um grande número de nossos informantes distancia o PGB do PB, pois, como vimos em Duarte (1993), no PB, a expressão *a gente* tem substituído o pronome *nós* e essa situação reduz o paradigma flexional do PB. Ainda sobre o PB, Omena (1996) analisa dados de pesquisas do uso de *a gente* em oposição a *nós*, que apontam maior frequência do primeiro em relação ao segundo, tanto para adultos quanto crianças em todas as funções gramaticais. Zilles (2007),

analisando dados de escrita, afirma que com o passar dos anos, nas sucessivas gerações, o grupo etário dos jovens tende a aumentar da expressão nominal *a gente* em detrimento do pronome *nós*.

Já no PE¹⁸, Duarte (1995) alega que não é muito frequente o uso da expressão *a gente* com referência definida. A autora encontrou pouquíssimas ocorrências e elas estavam ligadas a falantes com pouca escolaridade. Vianna e Lopes (2011) analisaram o uso de *a gente* nas variedades brasileira e europeia e constataram que a variação entre as formas de primeira pessoa do plural no PE, embora apresente semelhanças com o que se observa para o fenômeno no PB, é fundamentalmente determinada por fatores sociais e não caracteriza um processo de mudança em curso.

Quanto à terceira pessoa do discurso, como podemos observar no quadro 16, temos percentuais maiores em relação às ocorrências de sujeito nulo tanto para o singular com 68% (169 ocorrências) quanto para o plural com 68% (32 ocorrências). Percebemos que, na maioria das vezes, a identificação do sujeito nulo vai depender do sintagma nominal que lhe dá referência no discurso: quando não há um distanciamento desse sintagma, o sujeito pronominal tende a ser nulo, mas quando o referente fica distante no discurso, favorece o sujeito pleno, a exemplo, temos a sentença (50) abaixo:

(50) “A Guiné Bissau é bom pra morar, **pro** tem a cultura, **pro** tem várias etnias, **ela** também é um país laico”. (I4, GB, F, PF)

A sentença em (50) é interessante porque revela o uso de nulo do PE e do PB. Por ser um dado do último período, percebemos uma interferência do PB sobre o PGB, visto que no PE, se não há mudança do referente no discurso, os sujeitos são nulos, ou seja, se não há leitura contrastiva com outros referentes, há sempre a posição de sujeito vazia no PE. Enquanto que o PB é insensível a essas restrições.

Conforme o quadro 18, vemos que não houve ocorrências de segunda pessoa do discurso, independentemente do número. Há apenas duas ocorrências, uma com sujeito nulo e outra com sujeito pleno, com referência de segunda pessoa, porém, com especificação

¹⁸ Para uma análise mais detalhada sobre o uso da expressão pronominal *a gente* em PE, consultar *Gramática comparada de a gente – variação no PE*. Dissertação de mestrado, de Sandra Pereira (2003), realizado em Portugal.

gramatical de terceira pessoa, que é pronome *você*. É válido frisarmos que não houve ocorrência do pronome *tu* pleno ou nulo. Na sentença com ocorrência do pronome pleno, o pronome *você* ocupa o lugar do *tu*, e, na sentença em que o pronome *você*, ocupando o lugar do *tu*, aparece nulo, referência de segunda pessoa e a concordância verbal faz-se com a desinência zero de terceira pessoa, que veremos no exemplo abaixo:

- (51) “**Você** dorme tarde, **pro** acorda muito cedo e o sistema de ensino aqui é muito diferente a do meu país” (I5, GB, M, PI)

Como vemos, o pronome *você* é usado com referência arbitrária sendo pleno na primeira oração e nulo na segunda. Conforme Duarte (1995), em PB é muito recorrente o uso do pronome *você* arbitrário sendo pleno. Em relação aos dados de escrita, não podemos afirmar que este seja um uso recorrente no PGB, visto que apenas um dos informantes fez esse uso. Contudo, intriga-nos o fato de este dado ter sido usado por um informante dos períodos iniciais, dificultando a possibilidade de esse uso ser uma influência do PB.

Por fim, reforçamos o fato de o nosso tipo textual ter sido uma narrativa e, como já dissemos, esse é um tipo de texto majoritariamente escrito em primeira ou terceira pessoa. Assim, atribuímos a ausência de ocorrências de segunda pessoa, singular e plural, a esse tipo de texto.

No que concerne à desinência verbal, acreditamos que ainda seja cedo para propor um paradigma de flexão verbal do português guineense, visto que não obtivemos uso da segunda do plural e há um uso muito restrito da segunda pessoa do singular. Ainda assim, tivemos uma percepção clara do uso da desinência verbal exclusiva de primeira pessoa do singular *-o*, da desinência verbal *-mos* de primeira pessoa do plural e da desinência verbal *-m* relativa à terceira pessoa do plural.

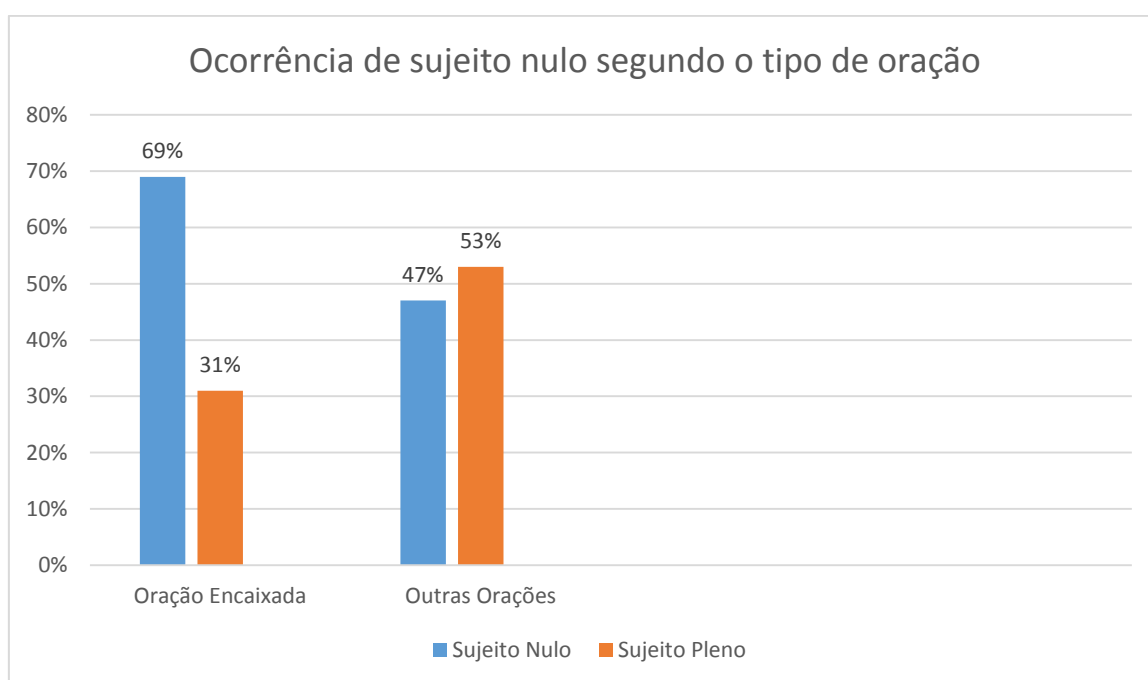
5.2.2 A estrutura da oração e seu estatuto sintático dentro do período

Assumimos com Duarte (1995) que as orações coordenadas com sequências de sujeitos com a mesma referência, que manifestam propriedade de elipse comuns a línguas que não admitem o sujeito nulo como uma propriedade sintática devem ser excluídas da análise,

visto que essas orações trariam falsas interpretações dos nossos dados por serem contextos obrigatórios de sujeitos nulos nas línguas naturais (SILVA, 2004). Sendo assim, agrupamos nossas orações em: oração encaixada e outras (que compreendem as orações independentes, as principais e as relativas).

Vejamos, no gráfico abaixo, a disposição percentual do tipo de oração em relação à ocorrência de sujeitos nulos:

Gráfico 9 - Ocorrência de sujeitos nulos segundo o tipo de oração (%)



Conforme o gráfico 9, as orações encaixadas apresentam um percentual considerável de sujeitos nulos (69%), em comparação aos outros tipos de orações que apresentam uma porcentagem maior para sujeitos plenos. De acordo com Duarte (1995), em línguas de sujeito nulo prototípico como o espanhol, o italiano e o PE, as orações encaixadas são contexto obrigatório de sujeito nulo. Segundo a autora, o uso de um sujeito pleno nessas orações acarretaria agramaticalidade da sentença. Já no caso do PB, temos que a dependência do referente que se cria a partir do contexto de uma oração encaixada facilita a ocorrência de sujeitos nulos, mesmo nessa língua que apresenta um sistema flexional enfraquecido. No português guineense, podemos observar a ocorrência de sujeito nulo em contexto de encaixadas nas sentenças abaixo:

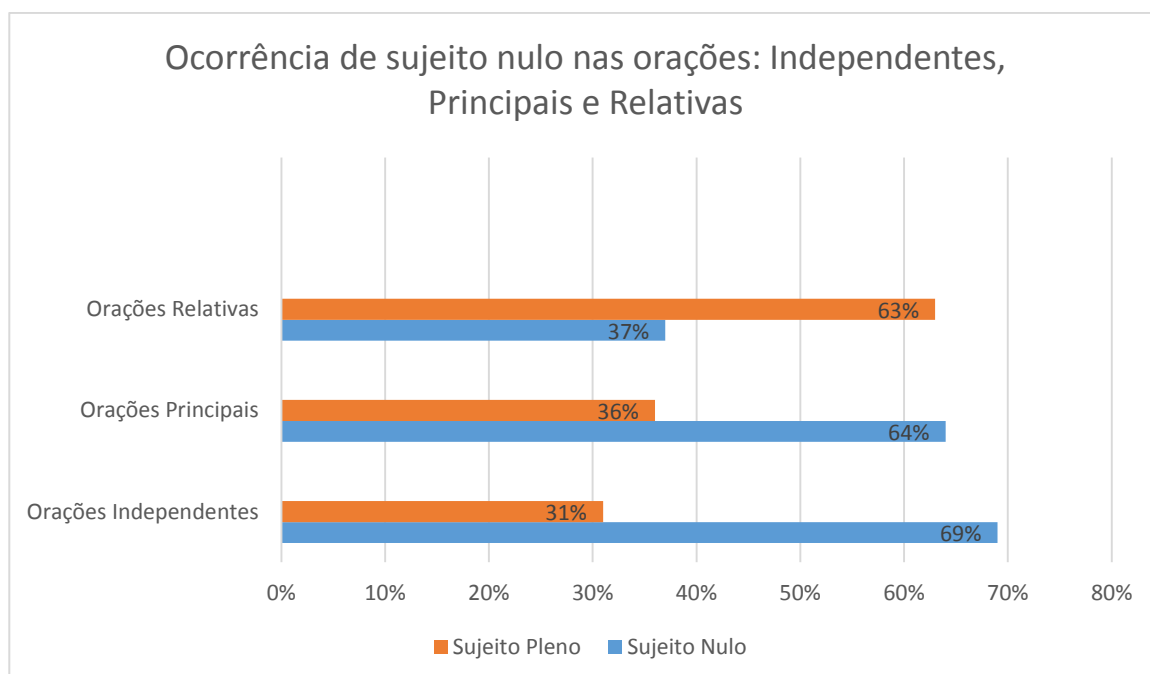
(51) “Espero que **pro** possa voltar logo” (I2, GB, F, PI)

(52) “Minha mãe diz que **pro** espera ansiosa pela minha volta” (I1, GB, M, PF)

Nossos resultados divergem dos resultados encontrados por Duarte (1995) para as ocorrências de sujeito nulo nas orações encaixadas. A autora encontrou um percentual de 32% para o PB, considerando um número muito baixo de ocorrências, visto que esse é um contexto de sujeito nulo obrigatório em línguas como o PE, o italiano e o espanhol.

Quanto às orações que foram classificadas como *outras*, fizemos um gráfico, tomando por base cada tipo de oração e a porcentagem de uso sujeitos nulos em cada uma delas, conforme apresentado a seguir:

Gráfico 10 - Ocorrência de sujeito nulo com base nas orações: independentes, principais e relativas (%)



No tocante às orações classificadas como *outras*, temos um percentual de 69% das ocorrências de sujeito nulo para as orações independentes, 64% para as principais e 37% para as orações relativas. Dentre essas orações, inclinamos atenção especial para as orações

relativas, pois essas apresentam um percentual de sujeito nulo mais baixo (37%) que o de sujeito pleno (63%). De acordo com Duarte (1995), temos que considerar a hipótese de que o preenchimento de Spec CP é o condicionamento mais forte ao uso do sujeito pleno no PE e, como temos observado, pelo menos, nos dados de escrita, o PGB tende a aproximar-se de sua língua-alvo, o PE. Abaixo, verificamos uma sentença em que o sujeito nulo da oração raiz é pleno na relativa:

(53) “**pro** tenho sonhado em voltar para o lugar onde **eu** nasci” (I5, GB, M, PI)

Quanto às orações independentes e principais, observamos, conforme o gráfico 10, uma preferência pelo uso do sujeito nulo.

Ao contabilizamos os dados, observamos uma preferência pelas ocorrências de sujeito nulo no PGB, um percentual elevado se comparado ao PB, nas primeiras pessoas do singular e do plural, desde então, imaginamos que as ocorrências de sujeito nulo nas orações principais e independentes teriam um percentual consideravelmente alto, em relação aos resultados encontrados por Duarte (1995) para o PB.

Vejamos as sentenças (54) e (55) a seguir:

(54) “**pro** Sou de uma família numerosa, que vivia numa harmonia e me fazia sentir um super-homem”. (I1, GB, M, PI)

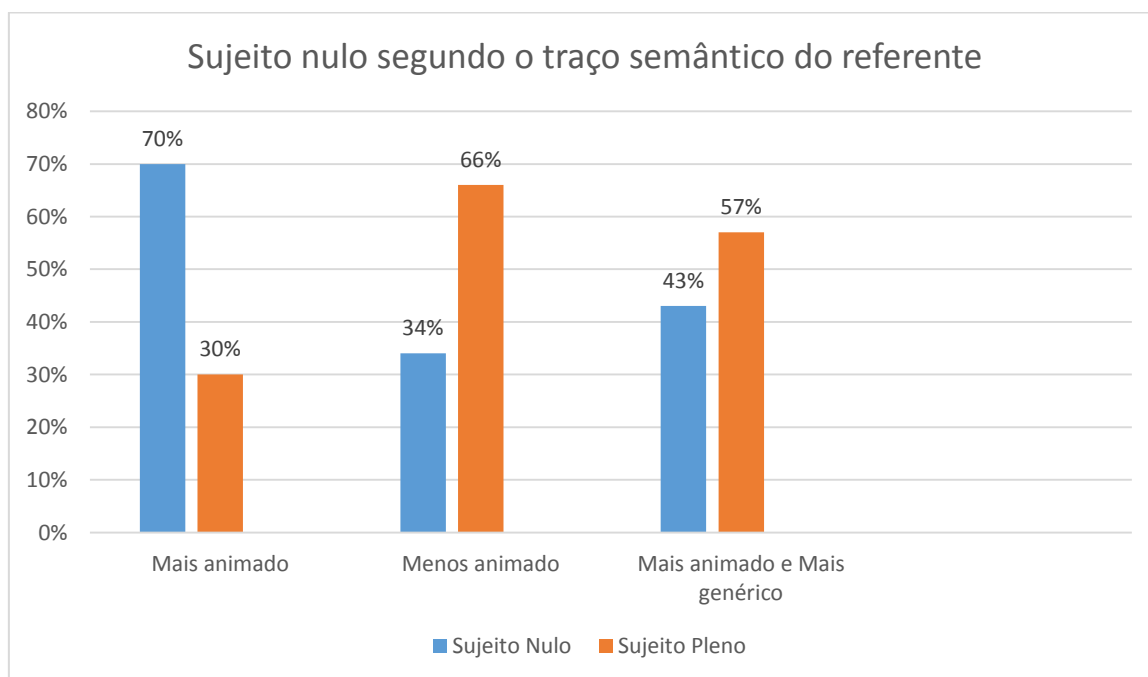
(55) “Primeiramente **pro** descrevo a minha terra natalino, como o mundo do paraíso da vida” (I2, GB, M, PI)

Como temos observado, até o presente momento, os resultados encontrados para o PGB, na modalidade escrita da língua parecem ir de encontro às nossas hipóteses, apresentando preferência pela ocorrência de sujeitos nulos em contextos que o PB apresenta preferencialmente sujeitos plenos. A seguir, analisamos o traço do referente de terceira pessoa.

5.2.3 O traço do referente de terceira pessoa

A partir de nossos estudos, constatamos que no PB o traço do referente de terceira pessoa é essencial para delimitar contextos de sujeito nulo, visto que sujeitos nulos de terceira pessoa são mais frequentes e estão submetidos a algumas restrições sintático-semânticas (cf. KATO e DUARTE, 2014). De acordo com Cyrino, Duarte e Kato (2000), o PB tem a opção de variantes nulas e não-nulas e, nesse caso, a hierarquia da referencialidade vai ser determinante na escolha de uma dessas variantes. Sendo assim, vejamos, no gráfico abaixo, quais foram os resultados do PGB para esta variável:

Gráfico 11 - As ocorrências de sujeito nulo com base no traço semântico do referente (%)



Observamos no gráfico acima que o referente [+] animado apresenta um percentual considerável de 70% de ocorrências de sujeito nulo, enquanto os referentes [-] animado e [+] animado e [+] genérico, as maiores ocorrências são de sujeitos plenos com percentuais de 66% e 57% de ocorrências de sujeitos plenos respectivamente. Esse resultado é contrário ao que foi encontrado por Duarte (1995), a autora deparou-se com 41% de sujeitos nulos para o traço [+] animado, enquanto temos 70%; para o traço [+] animado e [+] genérico, obteve 72% de ocorrências de sujeitos nulos, enquanto temos 34% e, para o traço [-] animado, encontrou 65% de nulos, enquanto temos 43%. Nas análises da autora o traço [-] animado representa um

contexto de resistência à implementação da mudança do parâmetro pro-drop para o não-pro-drop no PB.

Abaixo, podemos observar em (56), (57) e (58) exemplos de como as sentenças com os traços [+] animado, [+] animado e [+] genérico e [-] animado apareceram nos dados do português guineense, respectivamente:

(56) (referente: meu irmão) “**pro** foi quem sempre cuidou de mim, **pro** falou-me que não poderia perder essa chance, **pro** deu todo incentivo para chegar aqui”. (I3, GB, M, PI)

(57) (referente: minha família) “**pro** que me ajudou com o dinheiro da passagem, **eles** fizeram de tudo para eu realizar meu sonho” (I2, GB, F, PF)

(58) (referente: minha terra natal) “**pro** é o melhor lugar que conheci, **ela** é meu berço, **pro** é o lugar pra onde eu quero voltar” (I5, GB, M, PI)

Conforme Galves (1987), à medida que se perde a obrigatoriedade do uso de sujeito nulo em PB com o traço referente [+] animado, é possível atribuir a referência arbitrária a uma categoria vazia nessa posição. Como visto no gráfico 11 e exemplificado em (56), esse ainda não é o caso do PGB, ainda há nessa língua uma grande ocorrência de sujeitos nulos com o traço [+] animado, assim, não podemos apontar para uma interferência do PB no PGB, mas queremos salientar que pode haver mudanças nesses resultados ao analisarmos os dados da modalidade oral da língua, visto que há um monitoramento maior na modalidade escrita que motiva os dados escritos do PGB irem na direção da norma europeia do PE.

5.2.4 A duplicação do sujeito

A duplicação do sujeito acontece em estruturas frásicas em que o sintagma nominal sujeito é retomado por um pronome no domínio interno da frase com o qual mantém uma

relação de correferência. Segundo Duarte (1995), essas estruturas não são comuns em línguas pro-drop. No PB, por exemplo, essas construções são muito produtivas, segundo a autora, à semelhança do que ocorre no francês, uma língua não-pro-drop. Quanto à relação entre a estrutura do sujeito duplicado e o Parâmetro do Sujeito Nulo, temos que a entrada dessa estrutura no sistema linguístico de uma língua pro-drop pode leva-la à mudança do paradigma, Duarte (1995, p. 123) buscou relacionar “a entrada da estrutura no sistema à perda da propriedade pro-drop, que caracteriza línguas como o italiano e o espanhol e não admite o deslocamento à esquerda, que seria a própria negação do Princípio Evite Pronome”.

Diante do exposto, temos que nos dados escritos do PGB, porém, não encontramos um número representativo dessas construções, o que o aproxima do PE: das 270 ocorrências de sujeitos plenos, apenas duas apresentam construções de sujeito duplicado. São elas:

(59) ”**A minha mãe, ela** me incentivou muito, ela que fez minha inscrição” (I6, GB, M, PF)

(60) “**O meu país, ele** é muito pequeno e se encontra na Costa Ocidental da África” (I2, GB, F, PF)

Como observamos, as construções de duplicação sujeito que ocorreram nos dados escritos do PGB são de terceira pessoa, porém, conforme Silva, Ziober e Carvalho (2017), esse não é um uso comum às variedades africanas do português. Os autores fazem uma análise comparativa entre as variedades africanas, europeia e brasileira, observando as construções de sujeitos duplicados. Em relação ao PGB, concluem que nessa variedade prevalecem sujeitos duplicados de primeira pessoa em decorrência do contato com o crioulo guineense

Dessarte, compreendemos que, em relação ao quantitativo geral de ocorrência de sujeitos nulos e plenos de que dispomos, essas duas sentenças que apresentam sujeitos duplicados representam um número muito pequeno para afirmarmos que o PGB se distancia de uma língua pro-drop, mas salientamos que foram escritas por informantes dos últimos períodos, fato que, comedido, pode apresentar uma possível interferência do PB sobre o PGB, visto que esse tipo de estrutura é muito produtiva no PB (cf. DUARTE, 1995).

5.3 AS VARIÁVEIS LINGUÍSTICAS: SÃO TOMÉ E PRÍNCIPE

5.3.1 Pessoa gramatical e desinência verbal

Vejamos a frequência de sujeitos nulos encontrados em relação à pessoa gramatical, no quadro abaixo:

Quadro 19 - Ocorrência de sujeito nulo, na modalidade escrita do PST, segundo a pessoa gramatical

Pessoas do Discurso	Sujeito Nulo	Sujeito Pleno	Peso Relativo
1ª pessoa do singular	83% (194)	17% (40)	0,78
3ª pessoa do singular	61% (89)	39% (56)	0,52
1ª pessoa do plural	86% (19)	14% (3)	0,79
3ª pessoa do plural	58% (33)	42% (24)	0,55

Ao analisarmos o quadro acima, deparamo-nos com percentuais de ocorrências de sujeitos nulos mais altos que os percentuais de sujeitos plenos em todas as pessoas do discurso observadas, sobretudo, quando observada a primeira pessoa do singular (83%) e do plural (86%). A ausência de dados com a segunda, reiterando o que foi dito na seção anterior, pode ser atribuída ao tipo textual que escolhemos para coletar nossos dados.

Quando comparamos nossos resultados aos de Duarte (1995), que encontrou no PB um percentual de 36% e 29% de sujeitos nulos para a primeira pessoa do singular e do plural, respectivamente, verificamos que nossos resultados se afastam dos resultados encontrados pela autora. Assim, nosso percentual de ocorrências de sujeitos nulos, em primeira pessoa, aproxima-se do que foi encontrado pela autora para o PE, uma língua de sujeito nulo consistente. Contudo, destacamos que Duarte trabalhou com dados de fala, enquanto, até o momento, nossos resultados referem-se aos dados de escrita.

Observemos, a seguir, ora a presença de sujeito pleno ora a presença de sujeito nulo com primeira pessoa do singular (61) e do plural (62):

(61) a. “Graças a Deus **eu** tinha uma vida maravilhosa em São Tomé” (I1, STP, M, PI)

b. “Sempre **pro** tive boa educação e não corria nenhum risco de ser assaltado”

(62) a. “No meu país **pro** temos pessoas acolhedoras” (I3, STP, M, PI)

b. “**Nós** lá sabíamos de muitas coisas aqui do Brasil” (I5, STP, M, PI)

Diferentemente dos resultados encontrados no PGB, em relação à primeira pessoa do plural, não houve nenhuma produção da expressão pronominal *a gente* nos dados da modalidade escrita do PST. Essa observação leva-nos a pensar que o PB não interfere de fato na produção de sujeitos nulos nessas variedades africanas do português, pelo menos, até o presente momento, não encontramos grande relevância dessa interferência na modalidade escrita de dessas variedades. Como foi dito anteriormente, o uso da expressão pronominal *a gente* já substitui o pronome *nós* no português brasileiro e, portanto, reduz o paradigma flexional dessa língua. No entanto, no PE, *a gente* não é uma expressão recorrente (DUARTE, 1993, 1995), fato que faz aproximar o PST do PE.

No que concerne à terceira pessoa do discurso, temos mais ocorrências de sujeitos nulos em detrimento de ocorrências de sujeitos plenos, conforme os percentuais do quadro 19: são 58% para sujeitos nulos na terceira do plural e 61% de sujeitos nulos para a terceira pessoa do singular. Entendemos que a identificação do sujeito nulo de terceira pessoa vai depender do sintagma nominal que lhe dá referência, sendo assim, quando o referente permanece o mesmo para todas as orações, a posição de sujeito deve permanecer vazia, pois não contraste como novos referentes, à guisa do que acontece com o PE.

Dessa forma, em relação à ocorrência de sujeito nulo de terceira pessoa, temos que o uso do PST é similar ao uso do PE. Observemos o seguinte exemplo:

(63) “**A minha mãe** é o meu porto seguro, **pro** é quem faz tudo por mim, **pro** foi quem me incentivou a chegar aqui” (I4, STP, F, PI)

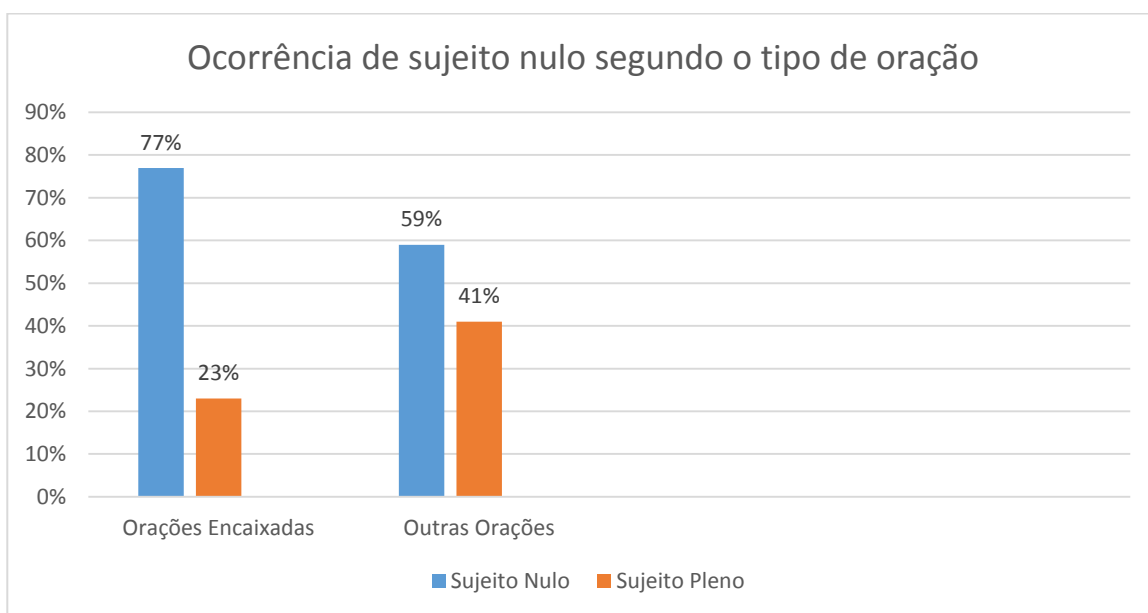
Sobre a desinência verbal, são usadas a desinência exclusiva de primeira pessoa do singular, a desinência *-mos* de primeira pessoa do plural e a desinência *-m* para a terceira

pessoa do plural. Contudo não podemos nos posicionar, até o momento, quanto a riqueza de flexão verbal do PST, visto que não obtivemos um número satisfatório de ocorrências em segunda pessoa para realizar tal análise.

5.3.2 A estrutura da oração e seu estatuto sintático dentro do período

Seguindo o mesmo viés de análise feito para o PGB sobre a variável tipo de oração, verifiquemos os resultados percentuais obtidos para o uso de sujeitos nulos no PST, tomando por base as orações encaixadas e outras orações:

Gráfico 12 - Ocorrência de sujeitos nulos no PST, segundo o tipo de oração (%)



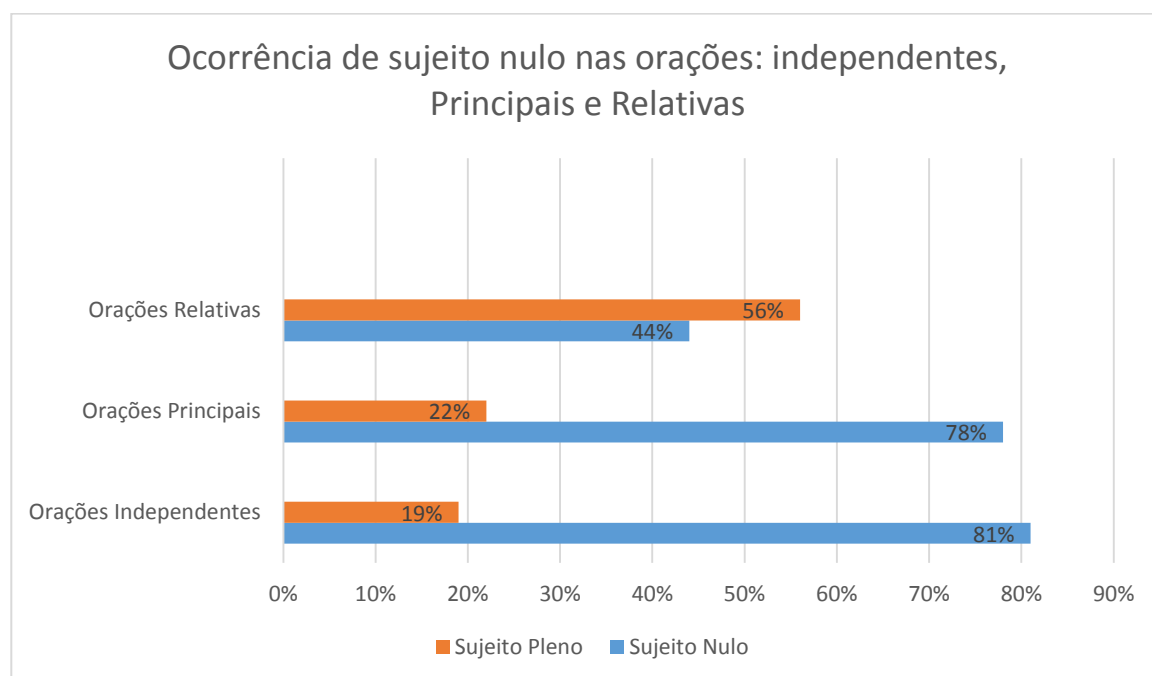
Consoante o gráfico acima, temos um valor de 77% de orações encaixadas com ocorrências de sujeitos nulos e 59% para as demais orações. Esses valores superam os que encontramos na análise dos dados do PGB. Ao que percebemos, até agora, o PST valida a nossa hipótese de ser uma língua mais próxima do PE, quando comparada ao PGB e ao PB. Vale lembrarmos que os são-tomenses, em sua maioria, já tem o PE como língua materna (cf. HAGEMEIGER, 2009) ao contrário dos guineenses que assumem o crioulo como língua materna. Observemos os exemplos, a seguir, que apresentam sujeitos nulos em contexto de orações encaixadas:

(64) “Meu Desejo é que **pro** acabe logo o curso e **pro** volte pra casa” (I2, STP, F, PI)

(65) “Recordo-me que **pro** sempre tive dificuldades em aprender” (I2, STP, F, PI)

Quanto às outras orações, analisamos as ocorrências de sujeito nulo e dispomos nossos resultados no gráfico abaixo:

Gráfico 13 - Ocorrência de sujeito nulo com base nas orações: independentes, principais e relativas (%)



No que diz respeito às orações acima, analisamos as orações independentes, as principais e as relativas que apresentaram um percentual de 81%, 78% e 44% de sujeitos nulos, respectivamente. Igualmente ao que foi visto nos dados do PGB (63%), as orações relativas apresentam maior percentual de sujeitos plenos no PST (56%) de ocorrências de sujeitos plenos. Conforme Duarte (1995), esse é um resultado recorrente também para o PE, visto que o preenchimento do Spec de CP é o condicionamento mais forte para o sujeito pleno nessa língua. Verificamos abaixo uma oração relativa com o sujeito pleno:

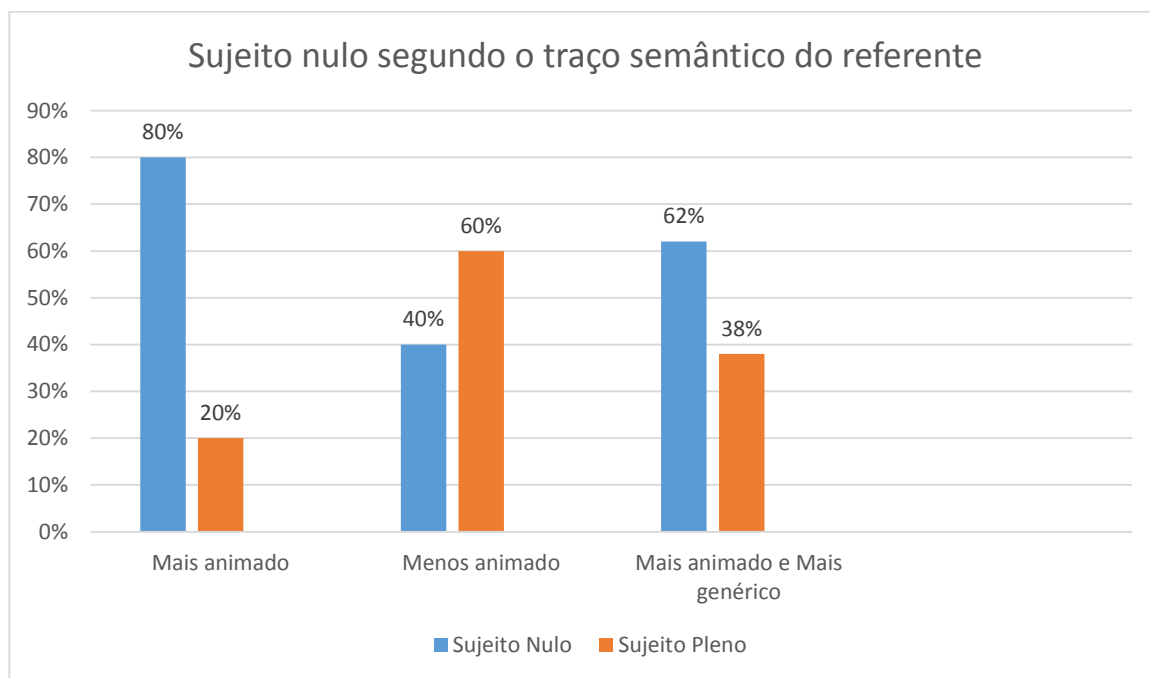
(66) “Meus amigos me ajudam com tudo que **pro** podem fazer por mim” (I4, STP, F, PF)

No que se refere às orações independentes e principais, encontramos um percentual consideravelmente alto para as ocorrências de sujeitos nulos com 81% e 78%. Como é sabido, trabalhamos com a hipótese de que o PST é uma língua que se aproxima do PE e, até agora, obtivemos resultados que corroboram essa hipótese, podendo ser essa variedade africana do português classificada como uma língua pro-drop consistente.

5.3.3 O traço do referente de terceira pessoa

Apresentamos no gráfico abaixo os resultados do PST em relação ao uso de sujeitos nulos, tomando por base o traço do referente de terceira pessoa:

Gráfico 14 - As ocorrências de sujeito nulo com base no traço semântico do referente (%)



O gráfico mostra-nos que o referente [+] animado apresenta 80% de ocorrências de sujeitos nulos, o [-] animado, 40% e o [+] animado e [+] genérico, 62%.

A produção de sujeitos nulos com referentes mais animados, nos dados do PST, é maior que a encontrada nos resultados do PGB (70%) e muito maior quando comparada aos resultados do PB (41%) obtidos por Duarte (1995). Observemos alguns exemplos de orações com os três tipos de referentes analisados:

(69) (referente: meu padraço) “**pro** trabalha como vigilante de uma igreja, **pro** cuida da casa e da minha família e **pro** é quem me ajuda financeiramente” (I3, STP, F, PI)

(70)¹⁹ (referente: meus amigos) “**pro** já estavam cá no Brasil quando **eu** cheguei, então foi mais fácil me adaptar, **eles** me ajudaram muito” (I2, GB, F, PF)

(71) (referente: São Tomé e Príncipe) “**pro** são duas ilhas maravilhosas, **pro** ficam no Golfo da Guiné e **pro** tem paisagens lindas” (I3, STP, F, PI)

Em suma, como observamos no gráfico 14, relativo aos dados escritos do PST, há um alto percentual de ocorrências de sujeito nulo para o referente [+] animado. Sendo assim, é possível concluirmos que não há interferência expressiva do PB sobre o PST e validamos nossa hipótese de que o PST se aproxima do PE.

¹⁹ O pronome *eu* é pleno porque contrasta com o primeiro referente já enunciado no domínio do discurso. E o pronome *eles* contrasta com o *eu*.

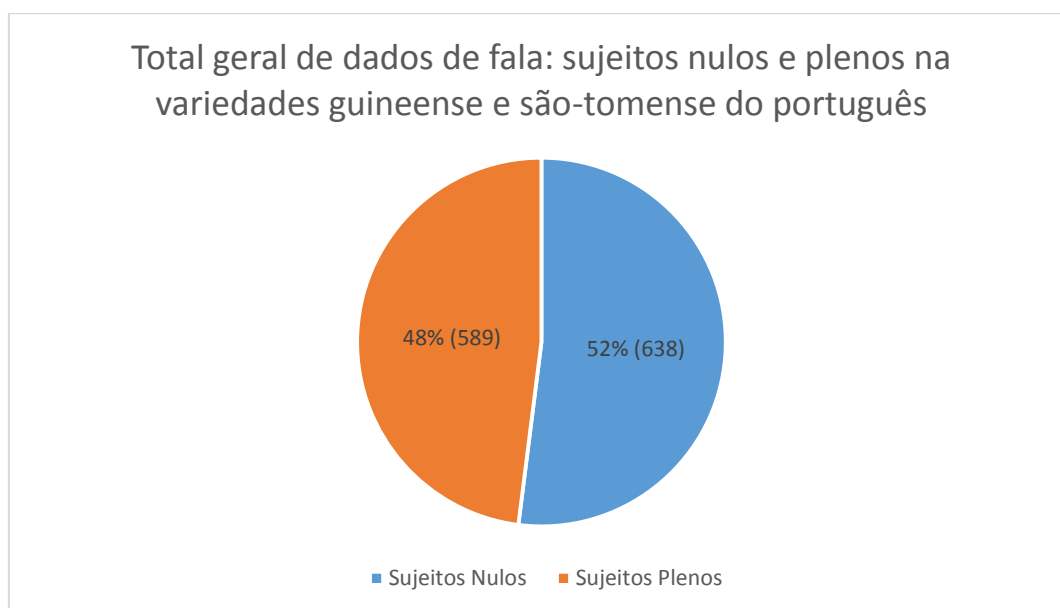
6 ANÁLISE DOS DADOS DE FALA

Este capítulo dará conta da análise de dados de fala. Na primeira subseção, faremos a comparação dos resultados quantitativos das duas variedades (PGB e PST) obtidos nas rodadas do programa Goldvarb X, conforme foi feito com os dados de escrita, e sintetizaremos as semelhanças e diferenças das variedades. Posteriormente, em uma segunda subseção, faremos a análise linguística dos dados de fala do PGB e, na terceira subseção, dos dados do PST.

6.1 COMPARANDO GUINÉ BISSAU E SÃO TOMÉ E PRÍNCIPE: QUANTITATIVO GERAL DOS DADOS

Contabilizamos, em nossos dados de fala, um total de 1227 sentenças com ocorrências de sujeitos nulos e plenos. Desse total, temos 642 sentenças para Guiné Bissau e 585 para São Tomé e Príncipe. De acordo com o total de dados que temos e considerando os sujeitos nulos e plenos, que é nossa variável dependente, obtivemos um resultado bem apertado, apresentando 52% (638 dados) para sujeitos nulos e 48% (589 dados) para sujeitos plenos. Vejamos o gráfico a seguir:

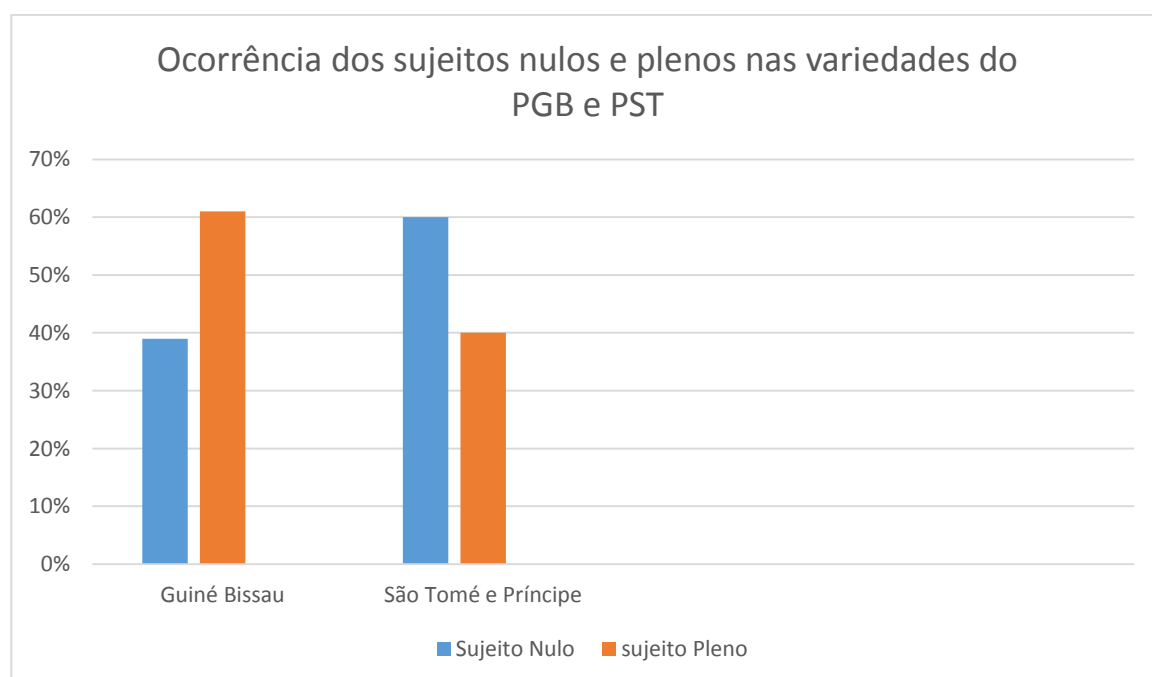
Gráfico 15 - Quantitativo geral e percentual de sujeitos nulos e plenos, nos dados de fala, do PGB e do PST



Ao compararmos os resultados obtidos entre os dados de fala e os de escrita temos um crescimento no percentual de sujeitos plenos que vai de 42% nos dados de escrita para 48% nos dados de fala. Esse número já pode nos indicar uma interferência mais massiva do PB sobre as variedades guineense e são-tomense do português. Vale ressaltarmos que, de acordo com nossos resultados, essa interferência pode ser melhor percebida nos dados de fala, visto que, no processo de escrita, o informante se monitora mais quanto ao uso da língua, principalmente, porque lidamos com informantes com maior escolaridade.

Quando comparados aos resultados de Duarte (1995) para o PE, temos que os dados de fala se distanciam da variedade europeia do português, visto que esta apresenta percentual de 66% para ocorrências de sujeitos nulos, enquanto encontramos um percentual de apenas 52% para sujeitos nulos em detrimento de 48% para sujeitos plenos. Vejamos no gráfico abaixo como esses dados distribuíram-se, tomando por base o país de origem dos alunos:

Gráfico 16 - Ocorrência de sujeitos nulos e plenos no PDG e no PST, tomando por base os dados de fala (%)



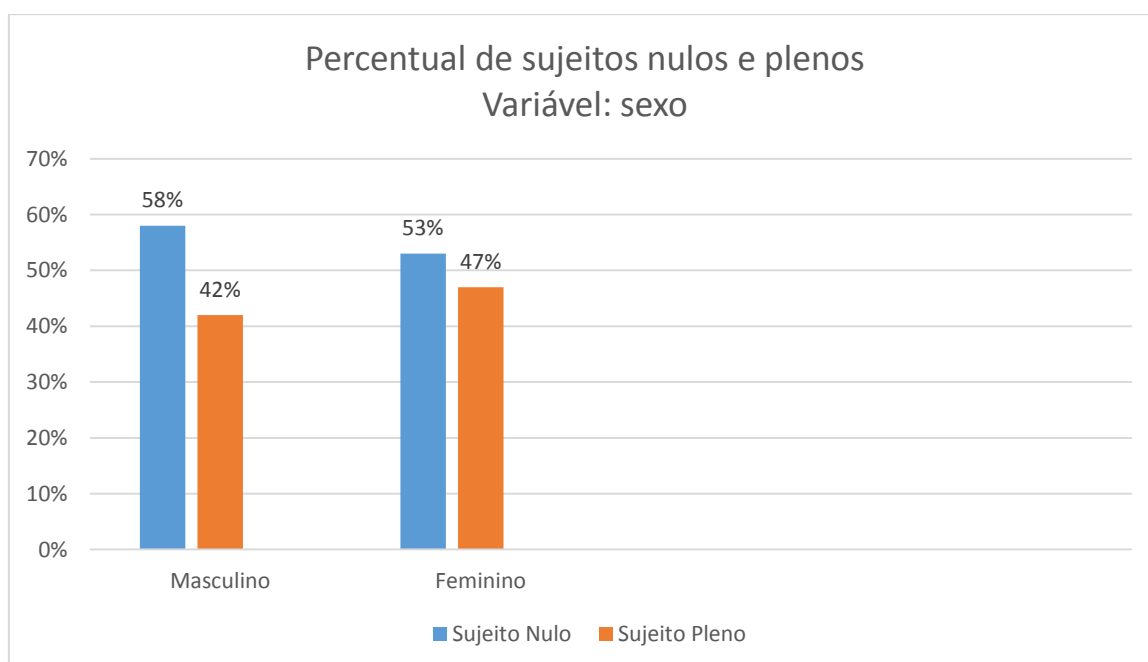
Em conformidade com o gráfico acima, vemos que os resultados são bem distintos quando comparados aos resultados dos dados de escrita. Houve um crescimento significativo das ocorrências de sujeitos plenos, da variedade guineense, quando confrontadas com as ocorrências desse mesmo sujeito obtidas nos dados de escrita, apresentando um percentual de

47% de sujeitos plenos nos dados de escrita e 61% nos dados de fala. No que concerne aos resultados da variedade são-tomense, temos um crescimento de apenas 4% de sujeitos plenos nos dados de fala, que passou de 36% nos dados de escrita, para 40% nos dados de fala. Esses resultados apontam para uma interferência mais efetiva do PB sobre os últimos períodos do PGB, dado que o percentual de sujeitos plenos (61%) apresentado por essa variedade destoa muito do que é encontrado para o PE (34%) por Duarte (1995). Enquanto, observamos que a diferença percentual em PST é muito pequena, tanto nos primeiros quanto nos últimos períodos.

Dentre as possíveis interferências do PB sobre o PGB e o PST, no que se refere ao uso do sujeito nulo, podemos citar: a preferência pelo preenchimento da posição de sujeito na maioria dos contextos, como ocorre em línguas não-pro-drop; o preenchimento da posição de sujeito com referencial [+] animado, contrariando o comum para línguas pro-drop e o enfraquecimento da flexão verbal. Essas são condições que podem levar ao uso do sujeito expresso, como vem sendo feito pelo PB. Por isso, investigamos em nossos resultados se há essa interferência do PB nos contextos de realização de sujeito nulo nas variedades africanas analisadas e o quanto isso modificou o uso dos falantes dessas variedades.

No que se refere à variável sexo, organizamos gráficos separados para analisar os resultados das duas variedades. Abaixo apresentamos os resultados para o PGB:

Gráfico 17 - Percentual de sujeitos nulos e plenos, realizados no PGB, tomando por base a variável sexo.



Como dissemos anteriormente, ao analisarmos os dados da modalidade escrita da língua, o uso de sujeitos nulos e plenos não é um fenômeno marcado socialmente, assim, os nossos resultados não podem ser atribuídos aos papéis sociais que nossos informantes assumem, de acordo com seu sexo. Conforme o gráfico 18, temos que os homens são responsáveis pela maior parte das ocorrências de sujeito nulo, apresentando 58% desses sujeitos, enquanto as mulheres apresentam um percentual de 54%.

A diferença percentual das ocorrências de sujeito nulo que encontramos na modalidade escrita da língua foi um pouco menor do que a encontrada nos dados de fala, tivemos 4% de diferença entre homens e mulheres nos dados de fala, enquanto tivemos apenas 1% de diferença nos dados de escrita. Abaixo identificamos alguns exemplos extraídos do *corpus* do PGB que evidenciam o uso de sujeitos plenos e nulos por homens e mulheres:

- *Sexo masculino:*

(72) a. “**Eu** prefiro falar crioulo, é a língua que **eu** falo desde cedo” (I2, GB, M, PI)

b. “**pro** Tenho muitas saudades de casa, da minha família” (I2, GB, M, PI)

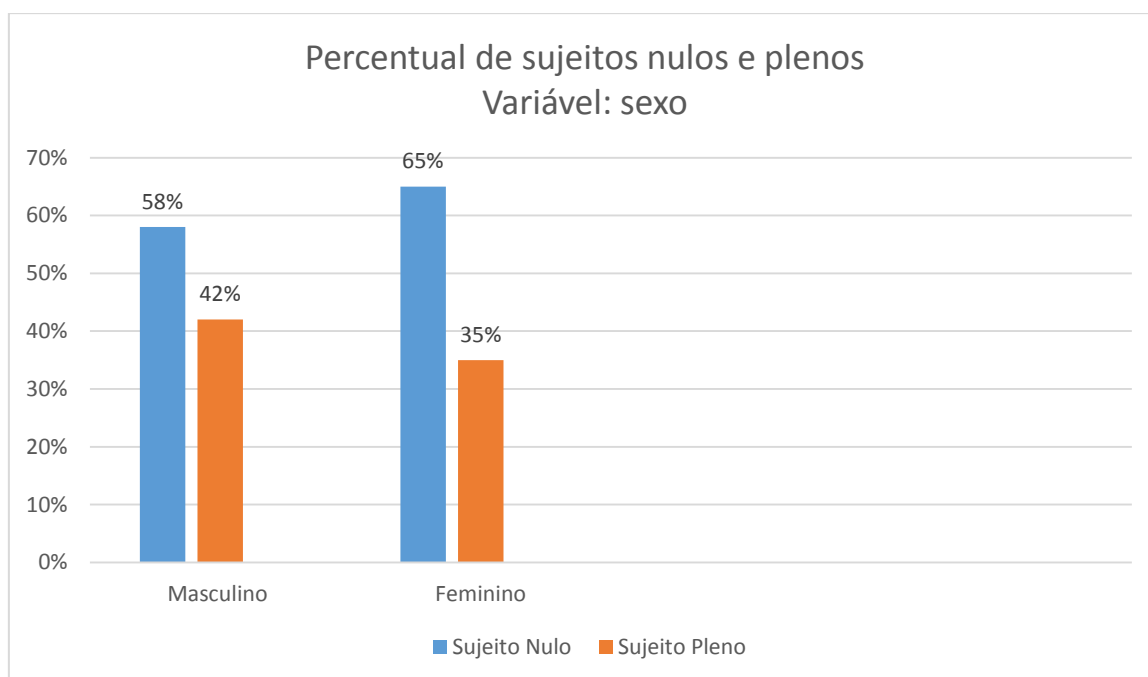
- *Sexo Feminino:*

(73) a. “**Eu** aprendi língua portuguesa na escola mesmo, **meus pais** falam crioulo”
(I4, GB, F, PF)

b. “**pro** Gosto de tapioca e de coxinha” (I4, GB, F, PF)

Em relação aos resultados de São Tomé e Príncipe, no que concerne à variável sexo, temos uma diferença maior, em relação a que foi encontrada no PGB entre as ocorrências de sujeitos nulos e plenos. Observemos no gráfico abaixo:

Gráfico 18 - Percentual de sujeitos nulos e plenos, realizados no PST, tomando por base a variável sexo.



De acordo com o gráfico 18, temos um percentual de 58% de ocorrências de sujeitos nulos para homens e 65% para mulheres, uma diferença mais expressiva que a encontrada nos dados de fala do PGB (6%) entre homens e mulheres. Como foi visto, no PST, as mulheres tendem a fazer uso de estruturas mais próximas à norma europeia do português, sempre pensando que essa é a língua de ascensão nesse país (cf. HAGEMEIJER 2009), contudo, acreditamos que a diferença percentual encontrada entre homens e mulheres (menos de 10%), seja uma diferença muito sutil para afirmar que o uso das mulheres é mais próximo ao PE.

Observamos abaixo alguns exemplos que evidenciam o uso de sujeitos nulos e plenos por homens e mulheres são-tomenses:

- *Sexo Masculino:*

(74) a. “**Eu** aprendi português em casa mesmo, com meus país” (I5, STP, M, PI)

b. “**pro** Quero voltar pra meu país e agradecer meu país por tudo ele fez por mim” (I5, STP, M, PI)

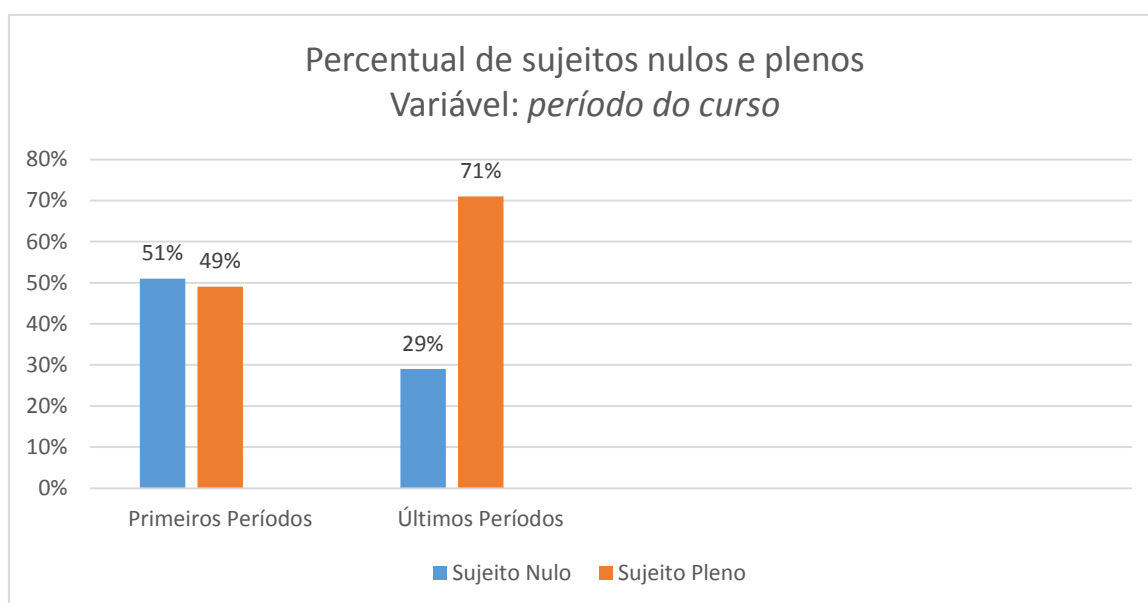
- *Sexo Feminino:*

(75) a. “**Eu** sofri preconceito aqui, no meu país não tem isso” (I2, STP, F, PI)

b. “Tudo no seu tempo, **pro** vou terminar meu curso e voltar pra casa” (I2, STP, F, PI)

Como observamos nos exemplos acima, homens e mulheres tem um comportamento muito próximo quanto ao uso da categoria vazia na posição sujeito. Sendo assim, a variável sexo não é grande significância para nossa análise independente da modalidade dos dados: escrita ou fala. Entretanto uma variável de grande peso para nossa análise é o tempo de permanência dos alunos no Brasil, pois, como já dissemos, julgamos que os alunos que estão nos períodos finais de seus cursos tendem a produzir mais sujeitos plenos que aqueles alunos recém-chegados ao Brasil, decorrente do contato com o PB. Vejamos no gráfico abaixo os resultados encontrados no PGB e no PST em relação ao período de tempo que os estudantes estão na universidade:

Gráfico 19 - Percentual de sujeitos nulos e plenos, realizados no PGB, tomando por base a variável período do curso.

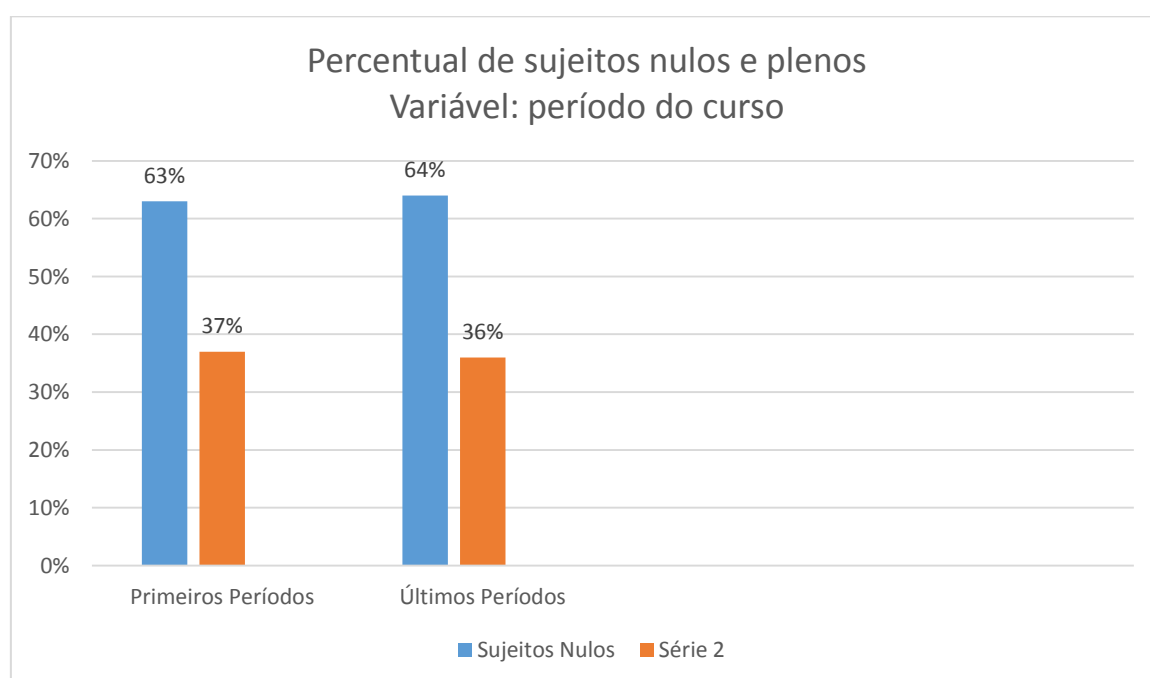


Conforme mostra o gráfico acima, os primeiros períodos apresentam resultados de 51% para sujeitos nulos e 49% para sujeitos plenos, sendo um resultado um pouco menor,

quando comparado aos dados de escrita, visto que chegamos à 53% para as ocorrências de sujeitos nulos. Já os últimos períodos, apresentam 71% para sujeitos plenos e 29% para sujeitos nulos, sendo o percentual de sujeitos plenos muito mais alto do que o encontrado nos dados de escrita (46%). Esse resultado confirma nossa hipótese de uma possível interferência do PB sobre o PGB, pois há um crescimento considerável no número de ocorrências de sujeitos foneticamente expressos por alunos que estão há mais tempo convivendo com o PB. Ainda sobre o PGB, cujo português é uma L2, ressaltamos novamente sobre a assimetria entre a língua falada (menos monitoramento) e a língua escrita (mais monitoramento). Na língua escrita, o grau de monitoramento é maior e o PB exerce pouca influência; já na fala, não só o grau de monitoramento é menor como o maior tempo de permanência no Brasil favorece mais plenos.

Sobre o PST, os resultados encontrados foram bem diferentes dos encontrados para o PGB quanto aos dados de fala. Porém, não se distanciaram dos resultados encontrados para o PST nos dados de escrita. Vejamos os percentuais no gráfico:

Gráfico 20 - Percentual de sujeitos nulos e plenos, realizados no PST, tomando por base a variável período do curso.



No gráfico 20, ao contrário do que aconteceu com o PGB, temos um aumento de 1% apenas nas ocorrências de sujeitos nulos, registrando 63% nos primeiros períodos e 64% nos

últimos. Quando comparamos esses resultados aos obtidos nos dados de escrita, percebemos um crescimento maior dos sujeitos plenos, visto que, nos dados de escrita, encontramos 38% nos primeiros períodos e 34% nos últimos (o crescimento de sujeitos nulos foi de 4%), enquanto, nos dados de fala, temos 37% de sujeitos plenos nos primeiros períodos e 36% nos últimos.

Como vemos, o aumento de sujeitos nulos não é tão grande nos dados de fala, dessa forma, entendemos que há uma pequena influência do PB sobre o PST, principalmente se comparada ao PGB. Acreditamos que esse fato aconteça em decorrência de o PST ser em geral a L1, assim, compreendemos que não depende apenas do grau de monitoramento da língua e do tempo de permanência no Brasil. Contudo, vemos que mesmo essa interferência sendo sutil do PB sobre o PST, ainda consegue travar um pouco o uso de sujeitos nulos pelos são-tomenses cuja língua materna é o PE, que é uma língua de sujeito nulo consistente. De acordo com Duarte (1995), um falante do PE com escolarização avançada, como são os nossos informantes, apresentam constantemente sujeitos nulos, sendo o uso de sujeitos plenos submetidos a restrições como contexto de contraste, de ênfase.

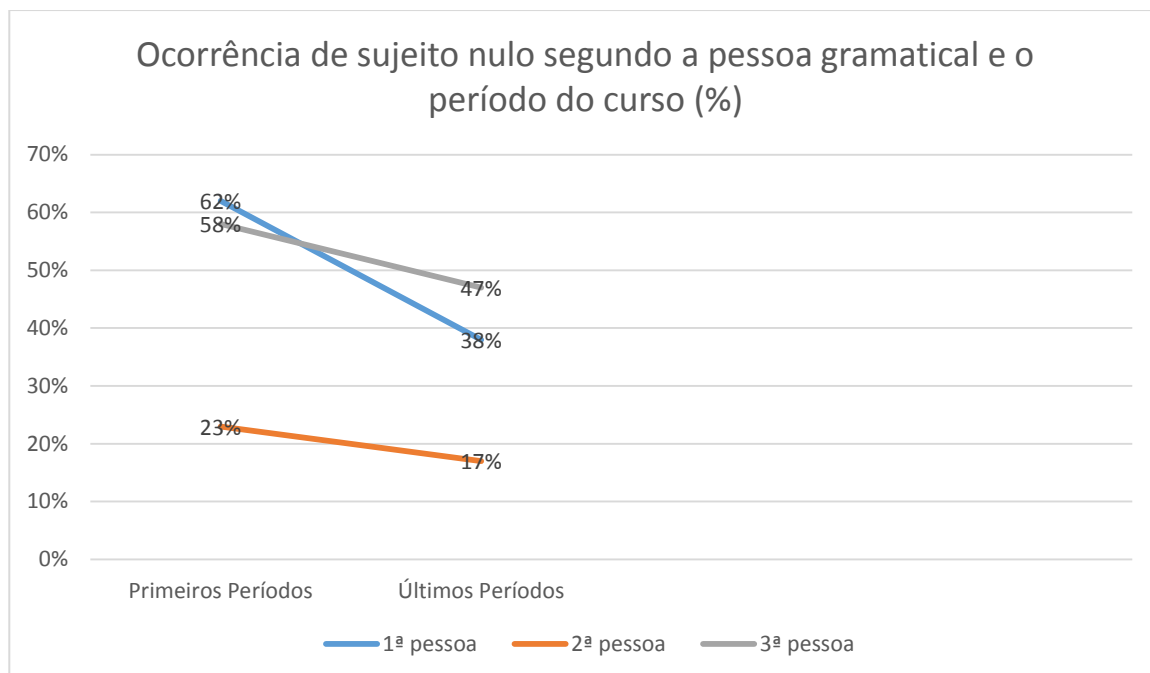
Em suma, os resultados encontrados nesta subseção mostram que o PTS é mais próximo ao PE, enquanto o PGB distancia-se mais dessa língua, corroborando nossa hipótese inicial. Em relação aos dados de fala e de escrita, constatamos a importância de fazer a análise nas duas modalidades da língua, pois encontramos resultados bem divergentes entre elas, principalmente no que se refere ao PGB. Ressaltamos ainda sobre a importância do status da aquisição do português para determinação do preenchimento ou não da posição de sujeito, visto que quando o português é adquirido como L1, no caso de São Tomé e Príncipe, houve uma menor interferência do PB sobre o PST. Em contrapartida, quando é adquirido como L2, como em Guiné Bissau, houve mais contextos de semelhanças com o PB.

6.2 AS VARIÁVEIS LINGUÍSTICAS: GUINÉ BISSAU

6.2.1 Pessoa gramatical e desinência verbal

Observemos a frequência de sujeitos nulos em relação à pessoa gramatical, tomando por base o tempo que os alunos vivem no Brasil (período do curso):

Gráfico 21 - Ocorrência de sujeito nulo, na língua falada do PGB, segundo a pessoa gramatical e o período do curso (%).



Analisando o gráfico, vemos que os menores percentuais de sujeitos nulos ocorrem na segunda pessoa, tanto nos primeiros (23%) quanto nos últimos períodos (17%). Em seguida, aparece a terceira pessoa do discurso com 58% da frequência de sujeitos nulos para os primeiros períodos e 47% para os últimos períodos. Por fim, e com índices mais altos que a terceira pessoa, vem a primeira pessoa do discurso apresentando 62% de sujeito nulo para os primeiros períodos e 38% para os últimos. Esses índices são mais baixos que os encontrados para os dados de escrita, mas ainda assim, são percentuais altos encontrados nos primeiros períodos, se comparados aos resultados do PB encontrados por Duarte (1995). Como bem sabemos, seguindo a hipótese da escala de referencialidade de Cyrino, Duarte e Kato (2000), as pessoas que teriam menos sujeitos nulos no PB seriam a primeira e a segunda. Porém, sabemos que no PE apenas a primeira e a segunda pessoa admitem sem restrições a omissão do sujeito nulo, pois a morfologia verbal destas permite a identificação com um referente sem ambiguidade (cf. DUARTE, 1995).

Como pudemos perceber, o gráfico 22 mostra uma queda percentual da frequência de sujeitos nulos em todas as pessoas dos primeiros períodos para os últimos, apresentando assim uma possível influência do PB na distribuição de sujeitos nulos na língua escrita dos alunos guineenses residentes no Brasil. Ao comparar nossos resultados com outros estudos, deparamo-nos com Oliveira e Santos (2007) que afirmam que em relação às primeiras pessoas

do singular e do plural, o português angolano apresenta uma característica que o aproxima do PB, visto que o preenchimento da posição de sujeito está ligado à hierarquia de referencialidade. Ainda sobre o português angolano, Teixeira (2012), afirma que, dentre as pessoas do discurso, a primeira pessoa do plural é a que apresenta o menor índice de sujeitos nulos, assim, aproximando-se do PB. Dessa forma, constatamos que o PGB, segundo nossa amostra, colhida com informantes residentes no Brasil, assim como o PA, apresenta indícios de estar caminhando em direção ao preenchimento da posição por sujeitos plenos, sobretudo, os de primeira pessoa, o que corrobora a hierarquia da referencialidade.

Quanto à observação do paradigma flexional, iniciamos com a ocorrência de sujeitos nulos de primeira pessoa, apresentada no quadro 20 abaixo:

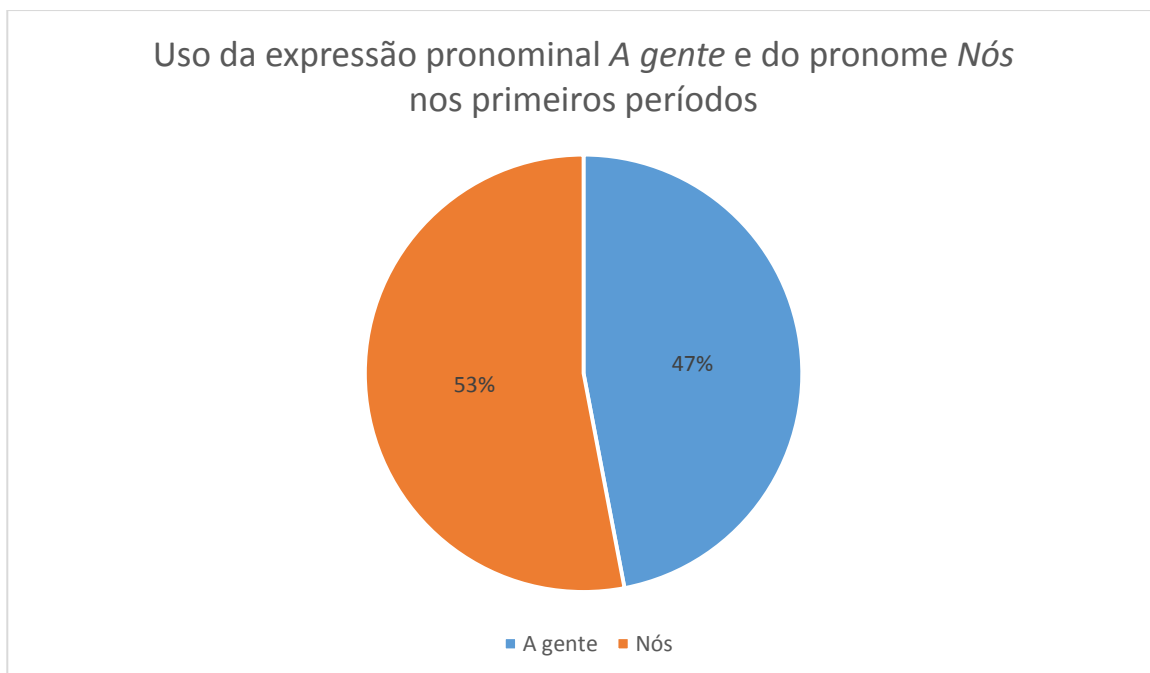
Quadro 20 - Ocorrência de sujeitos nulos de primeira pessoa no PGB, tomando por base os primeiros e últimos períodos (%)

Pessoa	Primeiros Períodos	Últimos Períodos	Peso Relativo
1ª pessoa do singular	60% / 70	49% / 86	0,57
1ª pessoa do plural	80% / 32	40% / 20	0,61

A partir da disposição dos dados no quadro 20, podemos perceber que, quando tratamos dos primeiros períodos, o percentual de nulos apresenta um índice alto na primeira pessoa, independentemente do número (60% e 80%), enquanto os últimos períodos apresentam índices mais baixos que os primeiros, independentemente de ser singular ou plural (49% e 40%). Assim, percebemos a influência do PB sobre o PGB nos últimos períodos.

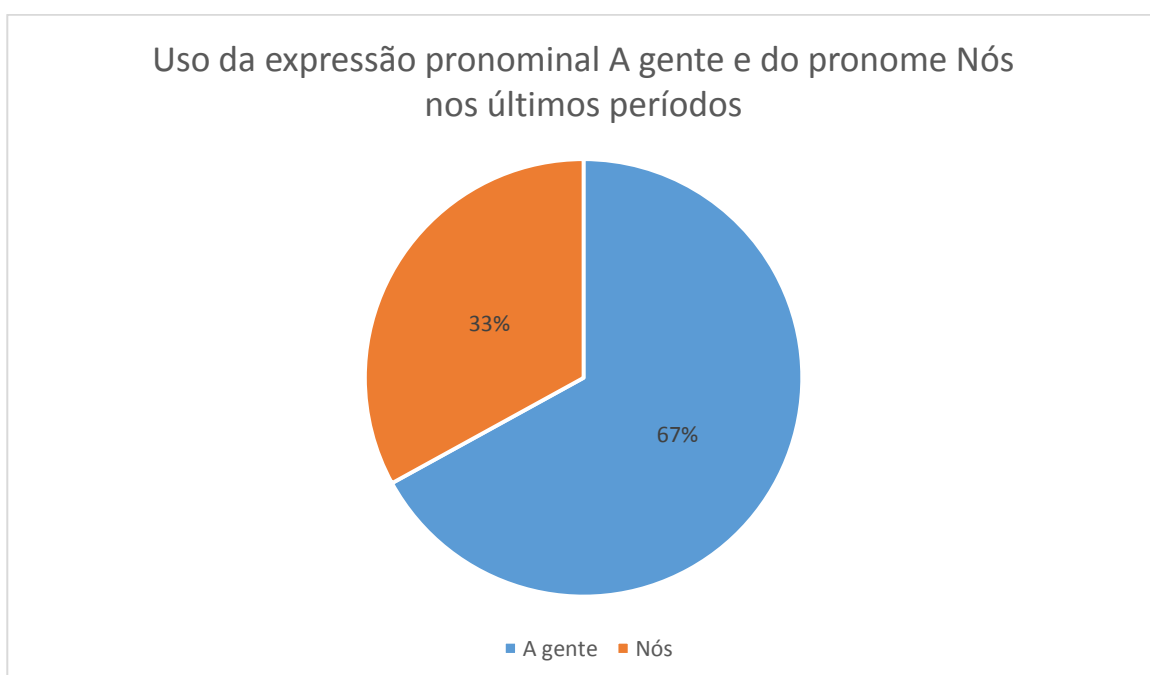
É importante dizer que, ao analisarmos a primeira pessoa do plural, constatamos que o uso do pronome *nós* aparece em quantidade (13 ocorrências) próxima à expressão *a gente* (19 ocorrências), no entanto são pouquíssimas as ocorrências do pronome *nós* como sujeitos plenos quando observados os primeiros períodos:

Gráfico 22 - Uso da expressão pronominal *A gente* e do pronome *Nós* plenos nos primeiros períodos do PGB



Por sua vez, nos últimos períodos, o uso da expressão *a gente* é superior ao uso do pronome *nós*, quando observamos as ocorrências de sujeito pleno:

Gráfico 23 - Uso da expressão pronominal *A gente* e do pronome *Nós* plenos, nos últimos períodos do PGB



Abaixo, nas sentenças (76) e (77)), podemos observar alguns exemplos do uso de sujeitos plenos com a expressão pronominal *a gente* nos primeiros e últimos períodos, respectivamente:

(76) “**A gente** vem pra cá cheio de esperança, em busca de um futuro melhor” (I4, GB, F, PI)

(77) “Já tinha algumas amigas aqui, então **a gente** foi morar na mesma casa e foi melhor pra se adaptar” (I3, GB, F, PF)

Quanto ao uso da expressão pronominal *a gente* não é um uso recorrente no PE, porém, no PB a expressão *a gente* tem substituído o pronome *nós* (cf. DUARTE, 1995). Em relação às outras variedades africanas do português, Teixeira (2012) afirma que especificamente em Luanda, capital de Angola, não se registra um uso expressivo de *a gente* mesmo entre falantes não alfabetizados, e, por isso, a concordância do PA não foi drasticamente enfraquecida. Acreditamos que o uso dessa expressão pronominal por falantes do PGB nos últimos períodos esteja diretamente ligado à influência do PB sobre essa língua, visto que no PE, língua-alvo para os guineenses, esse uso não é recorrente.

Em relação à distribuição de sujeitos nulos de segunda pessoa, observemos o seguinte quadro:

Quadro 21 - Ocorrência de sujeitos nulos de segunda pessoa no PGB(%).

Pessoa	Primeiros Períodos	Últimos Períodos	Total	Peso Relativo
2ª pessoa do singular	5 ocorrências (60%)	3 ocorrências (50%)	(57%)	0,52
2ª pessoa do plural	Não houve ocorrências	Não houve ocorrências	-	

Conforme verificado no quadro 21, os percentuais de sujeitos nulos de segunda pessoa são mínimos, assim como aconteceu nos dados de escrita, tanto nos primeiros quanto nos

últimos períodos. Não há ocorrências de sujeitos nulos na segunda pessoa do plural e há apenas 57% de sujeitos nulos com referência de segunda pessoa do singular, no total entre primeiros e últimos períodos.

Assim como nos dados de escrita, houve a ocorrência do *você* com referência arbitrária também nos dados de fala. As frases que continham sujeitos plenos e nulos com referência de segunda pessoa, porém com desinência gramatical de terceira pessoa do singular, em sua maioria, apresentavam uma interação com o entrevistador, como podemos em (78) e (79):

(78) “Quando **você** conhecer meu país, **você** vai ver, é lindo demais”. (I.4, GB, M, PF)

(79) “Não **pro** sabe como é ruim ter que se afastar de todo mundo para estudar”. (I.2, GB, F, PF)

Duarte (1995) afirma que esse uso é muito recorrente no PB, porém não é um uso comum no PE. Assim, atribuímos esse uso feito nos últimos períodos por falantes do PGB à influência exercida pelo PB:

(80) “**Você** chega aqui e é tudo muito novo, **você** fica meio perdido” (I2, GB, F, PF)

É válido frisarmos que não houve ocorrência do pronome *tu* pleno ou nulo tanto nos primeiros quanto nos últimos períodos. Nas sentenças com ocorrências do pronome pleno de segunda pessoa, o pronome *você* ocupa o lugar do *tu*. Isso mostra um certo enfraquecimento da morfologia de flexão verbal em PGB, ao contrário da língua falada no PE (cf. Duarte 1995; 2000), o que aproxima o PB do PGB. Esse dado também revela a influência do crioulo guineense cuja morfologia de flexão verbal é pobre. Dessa maneira, reforçando a hipótese da hierarquia referencial proposta por Cyrino, Kato, Duarte (2000). Observamos, então, que na língua falada do PGB, as pessoas que mais possuem ocorrências de sujeitos plenos são as primeiras e as segundas, o que não se observa em PE.

Por fim, no que concerne aos resultados quantitativos referentes aos sujeitos nulos de terceira pessoa, obtivemos o seguinte:

Quadro 22 - Ocorrência de sujeitos nulos de terceira pessoa nos dados de fala do PGB (%)

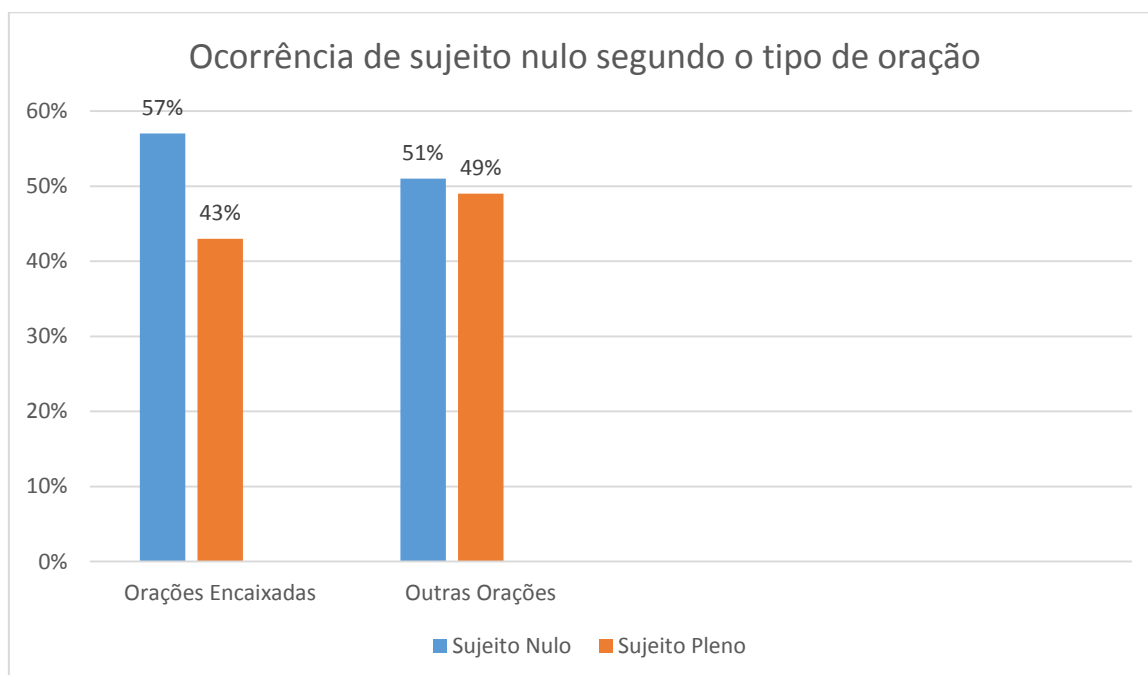
Pessoa	Primeiros Períodos	Últimos Períodos	Peso Relativo
3ª pessoa do singular	40%	58%	0,57
3ª pessoa do plural	54%	51%	0,53

De acordo com o quadro 3, de maneira geral, percebemos que há uma preferência por sujeitos nulos, tanto no singular (51%) quanto no plural (52%) da terceira pessoa, porém essa preferência não aparece com uma diferença significativa quando contrastada com a percentagem de sujeitos plenos. Conforme Duarte (1995), a terceira pessoa é um contexto de resistência ao uso do sujeito pleno também no PB.

Ao contrastarmos os dados dos primeiros períodos com os últimos, temos um crescimento de 18% da ocorrência de sujeitos nulos em terceira pessoa do singular dos primeiros períodos (40%) para os últimos (58%), em contrapartida, há um decréscimo de 3% nas ocorrências de sujeitos nulos em terceira pessoa do plural dos primeiros períodos (54%) para os últimos (51%). Assim, entendemos que o PGB parece seguir o PB quanto à escala de referencialidade, preferindo a posição de sujeito vazia para terceira pessoa do discurso.

6.2.2 A estrutura da oração e seu estatuto sintático dentro do período

Observemos, no gráfico abaixo, os resultados percentuais obtidos para os tipos de oração, em relação à ocorrência de sujeitos nulos:

Gráfico 24 - Ocorrência de sujeitos nulos nos dados de fala do PGB, segundo o tipo de oração (%)

Conforme vemos no gráfico 25, há um percentual de 57% de ocorrências sujeitos nulos nas orações encaixadas, enquanto, nas orações classificadas como *outras*, o percentual de sujeitos nulos (51%) é quase equivalente ao percentual de sujeitos plenos (49%). Em relação às orações encaixadas, ao compararmos os dados de fala com os dados de escrita, temos um decréscimo significativo, visto que, nos dados de escrita, o percentual de ocorrências de sujeitos nulos era de 69%. Acreditamos que esse decréscimo de 12%, que acontece dos primeiros para os últimos períodos, seja em decorrência do contato do PGB com o PB, pois, embora o contexto de encaixadas seja um contexto favorável ao uso de sujeitos nulos no PB, essa língua ainda apresenta índices de sujeitos nulos bem menores que o PE. Observemos um exemplo na sentença abaixo:

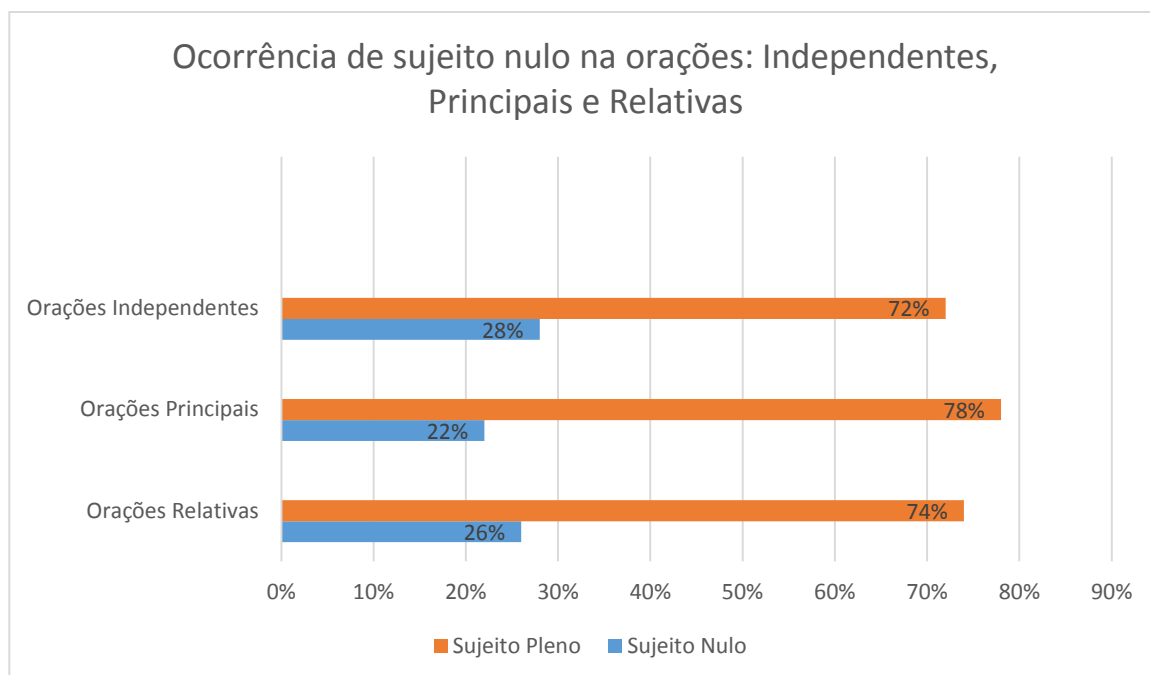
(81) “O importante é que **pro** termine meu curso” (I3, GB, F, PI)

Nossos resultados para os dados de fala do PGB aproximam-se mais do PB que os resultados obtidos para os dados de escrita. Podemos dizer que, em relação às orações encaixadas, o PGB encontra-se “no meio do caminho”, uma vez que apresenta um percentual

maior que o encontrado por Duarte (1995) para o PB (32%), mas apresenta um percentual baixo em relação ao que se espera de uma língua de sujeito nulo consistente:

Quanto às demais orações, vejamos os seguintes resultados no gráfico:

Gráfico 25 - Ocorrência de sujeito nulo em dados de fala do PGB com base nas orações: independentes, principais e relativas (%)



Conforme sabemos, as orações relativas apresentam um contexto favorável ao uso de sujeitos plenos até em línguas de sujeito nulo consistente como é o PE (cf. DUARTE, 1995). Encontramos, em nossos dados de fala, uma porcentagem ainda maior (72%) que a encontrada nos dados de escrita (63%). Esse resultado assemelha-se ao encontrado por Duarte (1993) para o PE e por Teixeira (2012) para o PA.

Quanto às orações independentes e principais houve uma redução muito grande das ocorrências de sujeito nulo nesses contextos, em relação aos resultados que obtivemos com os dados de escrita. Tínhamos um percentual de 64% de sujeitos nulos nas orações principais, nos dados de escrita, que passaram para 28%, nos dados de fala. As orações independentes apresentavam um percentual de 69%, em relação às ocorrências de sujeito nulo e passaram a apresentar 22% para os dados de fala. Relacionamos esse declínio nas ocorrências de sujeitos nulos, em contextos de orações independentes e principais ao crescimento de sujeitos plenos

nas primeiras pessoas do discurso. Observamos em nossas entrevistas que era muito comum que os informantes iniciassem suas respostas com a posição de sujeito preenchida, mesmo na primeira pessoa do discurso, que a flexão verbal daria conta de retomar o referente. Vejamos nas sentenças abaixo:

(82) “Eu estou aqui há quatro anos” (I3, GB, M, PF)

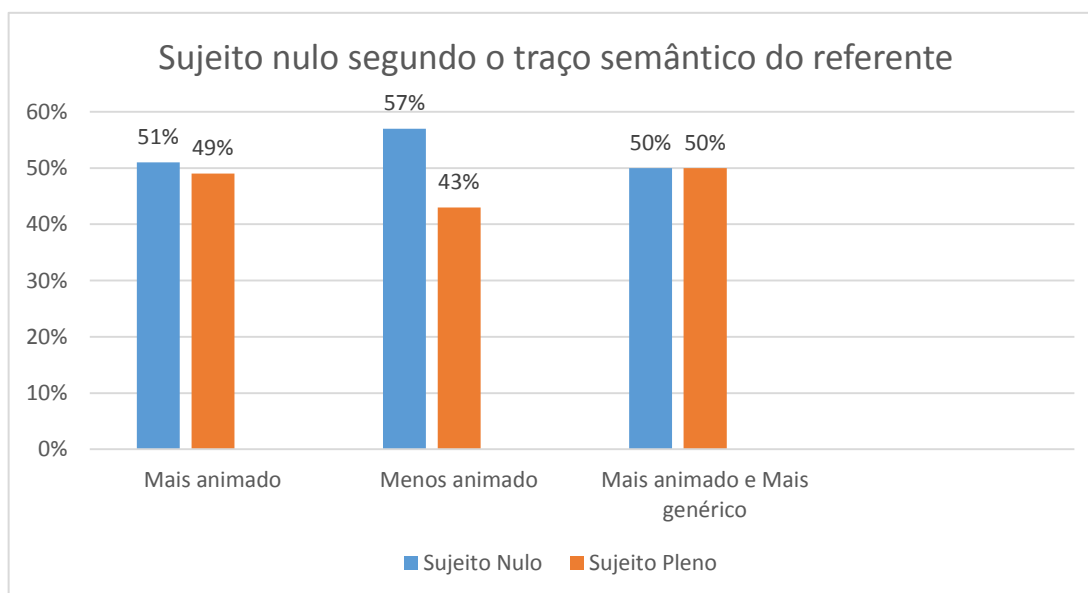
(83) “Eu tenho a necessidade de falar com minha mãe todo dia, ainda bem que existe a tecnologia” (I4, GB, F, PF)

Segundo nossos resultados, parece que a língua falada do PGB se aproxima muito do PB, pois essa língua licencia mais sujeitos plenos nas primeiras pessoas do discurso, faz uso recorrente da expressão nominal *a gente*, faz uso do *you* arbitrário, prefere os sujeitos plenos aos nulos nas orações independentes e principais e esses são contexto de sujeito nulo obrigatório em línguas pro-drop prototípica. Vimos que essa realidade não é a mesma quando tratamos dos dados de escrita, pois obtivemos resultados bem diferentes, assim, entendemos que o PB parece ter uma influência mais expressiva no PGB falado pelos nossos informantes.

6.2.3 O traço referente de terceira pessoa

Observemos o gráfico abaixo com nossos resultados:

Gráfico 26 - As ocorrências de sujeito nulo com base no traço semântico do referente (%)



O gráfico 26 apresenta-nos percentuais de 51% de nulos para o referente [+] animado, 57% para o referente [-] animado e 50% de ocorrências de sujeitos nulos para o referente [+] animado e [+] genérico. Esses resultados divergem dos que encontramos nos dados de fala, visto que tínhamos 70% de nulos para o referente [+] animado e 66% e 57% para o [-] animado e [+] animado e [+] genérico, respectivamente. Esse declínio das ocorrências de sujeitos nulos em relação ao traço [+] animado relaciona-se também com o aumento de sujeitos plenos nas primeiras pessoas do discurso e leva o PGB para mais próximo do PB. Nossos resultados dos dados de fala vão ao encontro dos resultados encontrados para o PB por Duarte (1995). Contudo, percebemos que os dados da autora, em relação ao traço [-] animado licencia mais sujeitos nulos que nossos dados, visto que Duarte (1995) obteve 65% de sujeitos nulos para esse referente.

Vejam nas sentenças abaixo como esse declínio nas ocorrências de sujeito nulo aconteceu:

(84) (referente: minha vó) “fala mandinga e crioulo, mas ela não fala português” (I3, GB, M, PF)

(85) (referente: minhas amigas) “estão me ajudando muito aqui, elas são uma família pra mim” (I5, GB, F, PI)

(86) (referente: minha língua) “é crioulo, foi ela que eu aprendi desde pequeno) (I4, GB, M, PF)

Como havíamos previsto na secção anterior, foi possível observar, nos dados de fala, a perca da obrigatoriedade do uso do sujeito nulo com traço [+] animado, aproximando o PGB do PB, no que se refere ao uso do sujeito pleno para tal referente.

6.2.4 A duplicidade do sujeito

Como dito na análise dos dados de escrita, não encontramos um quantitativo expressivo de ocorrências de sujeito duplo, nem mesmo nos dados de fala. Supomos que nesses dados haveria uma quantidade significativa dessas ocorrências, mas não foi isso que aconteceu. Encontramos apenas uma sentença com duplicidade de sujeito e foi de uma mesma informante que também usou esse contexto de sujeito pleno nos dados de escrita. Vemos essa construção em (87):

(87) "A minha terra, ela é maravilhosa" (I2, GB, F, PF)

Mais uma vez, não podemos afirmar que essa construção em específico represente um uso geral, principalmente por ter acontecido com um mesmo informante. Contudo, não podemos deixar de ressaltar que essa não é uma construção comum ao português europeu e que possivelmente seja reflexo da influência do PB no PGB.

6.3 AS VARIÁVEIS LINGUÍSTICAS: SÃO TOMÉ E PRÍNCIPE

6.3.1 Pessoa gramatical

Os resultados dos dados de fala de São Tomé e Príncipe, em geral, não diferiram muito dos resultados obtidos para os dados de escrita. Ao que nos parece, os são-tomenses são mais rigorosos com a norma culta da língua. Em nossas entrevistas, ouvimos de vários informantes são-tomenses que "o brasileiro não sabe falar o português, falam errado até na universidade" (I2, STP, F, PI). Diante dessa afirmação, podemos considerar que há, por parte dos são-tomenses, uma avaliação, a certo modo, negativa sobre a variedade brasileira quando a contrapõe à variedade europeia, que, conforme Hagemeijer (2009) é a variedade ensinada em São Tomé e Príncipe.

A partir dessas afirmações, seguimos para a análise dos dados de fala. Vejamos a frequência de sujeitos nulos encontrados em relação à pessoa gramatical, no quadro abaixo:

Quadro 23 - Ocorrência de sujeito nulo, na modalidade falada da língua, na variedade são-tomense, segundo a pessoa gramatical (%)

Pessoas do Discurso	Sujeito Nulo	Sujeito Pleno	Peso Relativo
1ª pessoa do singular	80%	20%	0,70
2ª pessoa do singular	(1 ocorrência)	0	-
3ª pessoa do singular	59%	41%	0,52
1ª pessoa do plural	80%	20%	0,68
3ª pessoa do plural	56%	44%	0,52

Em relação a pessoa gramatical, na variedade são-tomense do português, não encontramos diferenças expressivas entre os dados de fala e escrita, como foi encontrado na variedade guineense, conforme podemos observar no quadro 23.

Os percentuais são bem próximos ao que foi encontrado para os dados de escrita, variam em 1% ou 2% no crescimento do uso de sujeitos plenos. Mas, há uma ocorrência especificamente que nos chama atenção, há uma sentença com a segunda pessoa do singular com a posição de sujeito vazia. Vejamos em (88):

(88) “*pro* Entendes o que eu falo?” (I3, STP, F, PI)

Essa foi a única ocorrência no português são-tomense com o uso da segunda pessoa do discurso. Ao que podemos perceber, a informante usa a desinência número-pessoal -s de acordo com o paradigma flexional do PE, conforme quadro 3. Contudo, não podemos afirmar que esse é um uso recorrente no PST, visto que obtivemos apenas uma sentença com tal ocorrência.

Nossos resultados são muito próximos aos encontrados por Duarte (1993) para o PE e inversamente proporcional aos encontrados por Duarte (1995) para o PB, pois, conforme a autora, há no PB a preferência pelo sujeito expreso na primeira pessoa do discurso (singular e plural).

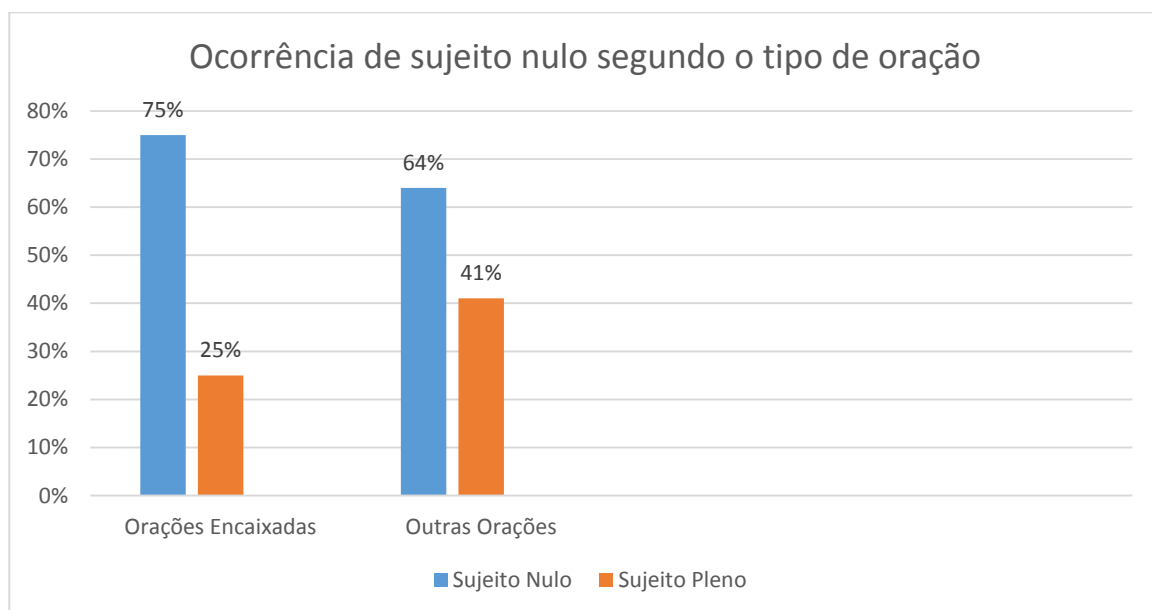
Em divergência com o PGB, mas de acordo com os dados de escrita do PST, em relação à primeira pessoa do plural, não houve nenhuma produção da expressão pronominal *a*

gente, nos dados da modalidade falada do português são-tomense, nem com referência definida tampouco com referência arbitrária. É um resultado que nos chama atenção, pois como a expressão *a gente* é muito difundida no português brasileiro, imaginamos que, ao menos nos dados de fala, ela seria uma expressão usada pelos são-tomenses, mas não encontramos essa ocorrência em nossos dados.

No que concerne à terceira pessoa do discurso, temos mais ocorrências de sujeitos nulos em detrimento de ocorrências de sujeitos plenos, seguindo o que foi visto também nos dados de escrita. Quanto à desinência verbal, pudemos observar apenas em uma sentença como a flexão verbal foi usada para a segunda pessoa do discurso, pois, como foi dito, obtivemos apenas uma sentença nesse contexto, mesmo assim, não podemos generalizar esse uso para os demais falantes, visto que apenas um fez esse uso. A desinência verbal exclusiva de primeira pessoa do singular e a desinência *-mos* de primeira pessoa do plural, como era esperado, são devidamente usadas e a desinência *-m* de terceira pessoa do plural também é sempre recorrente. Assim, inferimos que o paradigma flexional do português são-tomense é aparentemente rico e aproxima-se do português europeu.

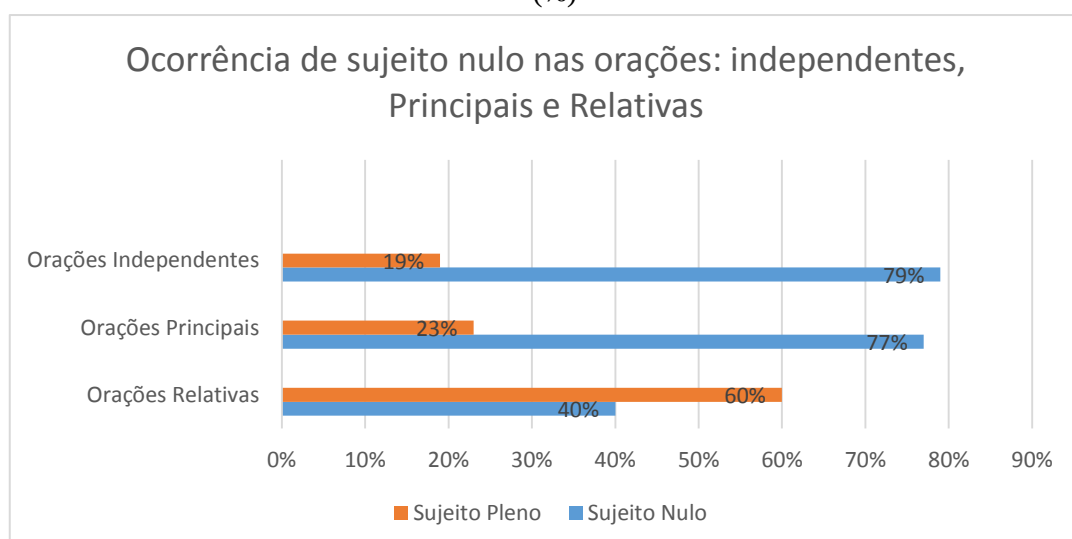
6.3.2 A estrutura da oração e seu estatuto sintático dentro do período

Vejamos no gráfico 12 a disposição percentual das orações, no que concerne à ocorrência de sujeitos nulos:

Gráfico 27 - ocorrência de sujeitos nulos segundo o tipo de oração (%)

De acordo com o gráfico acima, obtivemos 75% de sujeitos nulos nas orações encaixadas, esse número é 2% menor do que encontramos nos dados de escrita da variedade são-tomense e é um resultado bem mais elevado quando comparado ao percentual de sujeitos nulos encontrados nos dados de fala de Guiné Bissau. Ao que percebemos, até agora, o português são-tomense valida nossa hipótese de ser uma língua mais próxima do português europeu, quando comparada ao PGB e ao PB.

Quanto às outras orações, analisamos as ocorrências de sujeito nulo e dispomos nossos resultados no gráfico abaixo:

Gráfico 28 - Ocorrência de sujeito nulo com base nas orações: independentes, principais e relativas (%)

No que diz respeito às orações agrupadas como outras, analisamos as orações independentes, as principais e as relativas que apresentaram um percentual de 79%, 77% e 60% de sujeitos nulos em cada uma delas, respectivamente. Dentre esses dados, gostaríamos de sinalizar as relativas que aumentaram consideravelmente seu número de sujeitos plenos, mas como já vimos, esse é um contexto comum para sujeitos plenos até em línguas de sujeito consistente como o PE. Ressaltamos que esses são resultados muito próximos aos que foram obtidos para os dados de escrita da variedade são-tomense, mas distanciam-se bastante quando comparados aos resultados obtidos para os dados de fala da variedade guineense.

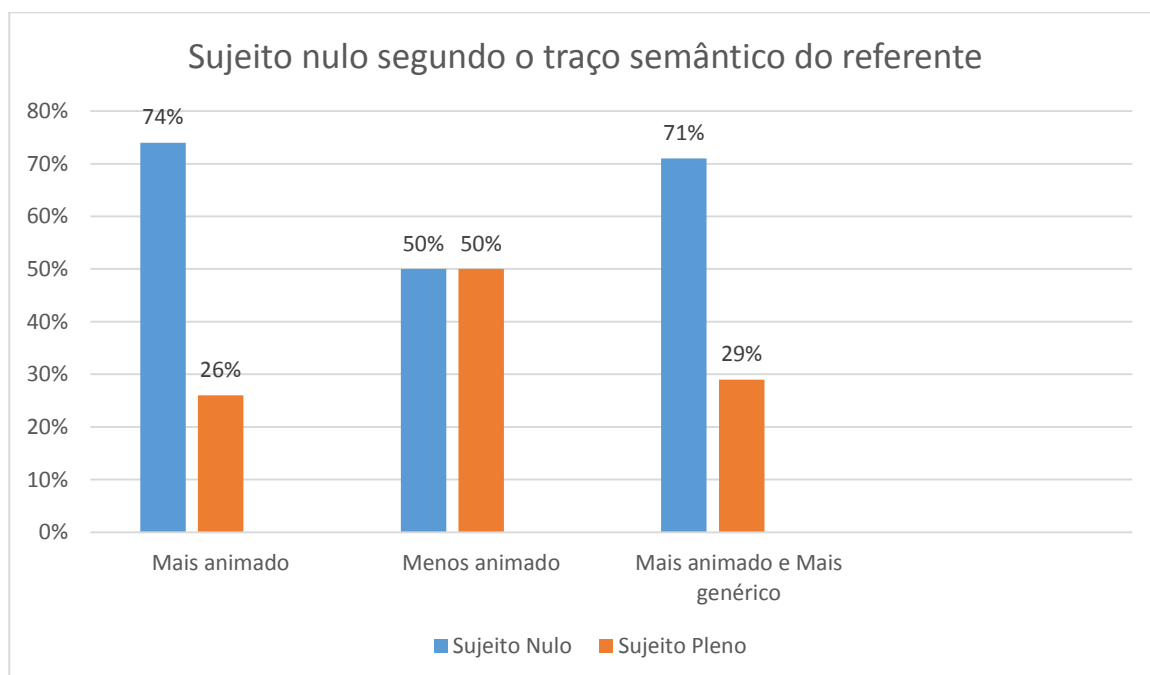
Mais uma vez, as conclusões que chegamos quanto ao PST leva-nos ao encontro das conclusões obtidas por Duarte (1993) para o PE e distanciam essa variedade africana do PB ou da influência que supomos que essa língua poderia exercer sobre o PST.

No que se refere às orações independentes e principais, ainda nos deparamos com altos contextos de produção de sujeitos nulos, mesmo tendo uma redução de 2% dessas ocorrências, em relação aos resultados encontrados para os dados de fala. A partir de tais resultados, validamos nossa hipótese de que o PST é uma língua que apresenta contextos de produção do sujeito nulo que são semelhantes aos contextos de produção de sua língua-alvo, o português europeu.

Os percentuais de ocorrências de sujeitos nulos encontrados nos mais diversos contextos do PST correspondem aos encontrados por Duarte (1993) para o português europeu, assim, reafirmamos nossa hipótese de que o português são-tomense se aproxima do português europeu no que diz respeito aos contextos de produção do sujeito nulo.

6.3.3 O traço do referente de terceira pessoa

Apresentamos no gráfico abaixo os resultados do PST em relação ao traço do referente de terceira pessoa, como dito na secção anterior, como sabemos, no PB esse é um contexto que freia a produção de sujeitos plenos. Pelo que vimos nos dados de escrita, o PST também vai ao encontro do PE nesse contexto. Abaixo, dispomos nossos resultados para analisar as ocorrências de sujeitos nulos em relação à variável em questão:

Gráfico 29 - As ocorrências de sujeito nulo com base no traço semântico do referente (%)

O gráfico acima mostra-nos percentuais um pouco menores, quando confrontados com os dados de escrita, houve uma redução de 2% e 3% nos referentes [+] animado e [+] animado e [+] genérico em relação às ocorrências de sujeito nulo, porém esses percentuais ainda maiores se comparados aos resultados da variedade guineense do português. Essa produção de sujeitos nulos no PST, em relação ao referente, é inversamente proporcional aos resultados encontrados para o PB, essa língua apresenta no traço [-] animado um contexto favorável à produção de sujeitos nulos, porém apresenta índices baixos de sujeitos nulos no traço [+] animado e esse é um contexto essencial para produção de sujeitos nulos em línguas pro-drop prototípica.

Dessa forma, ratificamos que o PST é uma língua que vai ao encontro da variedade europeia do português e é favorável à produção de sujeitos nulos, mesmo em contextos em que o PB teria preferência por sujeitos plenos.

7 CONSIDERAÇÕES FINAIS

Neste estudo, investigamos a interferência do português brasileiro na língua falada e escrita de alunos guineenses e santomenses dos primeiros e últimos períodos da UNILAB-CE no que concerne à distribuição de sujeitos nulos e plenos.

Pelo que vimos, os falantes guineenses e são-tomenses adquirem a língua portuguesa de maneiras distintas, enquanto aqueles adquirem como L2, estes adquirem como L1, conforme Couto e Embaló (2010) e Hagemeijer (2007; 2009).

Com base em um corpus oral e escrito resultante da coleta de 48 textos narrativos e 48 entrevistas informais cujos dados foram submetidos ao programa computacional GoldVab X, evidenciamos que o PB influencia mais expressivamente no PGB, enquanto que no PST são encontradas apenas algumas situações que mostram essa influência muito sutilmente.

Centrando nossa atenção nesse momento aos dados de fala, constatamos que o PGB e o PST encontram-se a meio caminho entre o PB e o PE, pois apresentam um percentual de sujeitos nulos (58%) maior que o encontrado para o PB (26%) e menor que o encontrado para o PE (66%). Quando analisadas individualmente, o PGB apresentou 53% de ocorrências de sujeitos nulos e o PST 64%. Esses resultados corroboraram nossa hipótese de que estudantes guineenses produziram mais sujeitos plenos em decorrência do uso frequente de sua língua materna, o crioulo guineense, uma língua não pro-drop, ao contrário do que foi verificado em São Tomé e Príncipe, em que a maioria da população já tem o português, uma língua de sujeito nulo consistente, como primeira língua (HAGEMEIJER, 2009).

No que se refere à variável sexo, os resultados de Guiné Bissau mostraram que 48% das mulheres realizaram o sujeito nulo, enquanto os homens, 52%, sendo, portanto, mínima a diferença percentual entre eles. Já os resultados de São Tomé e Príncipe apresentaram 68% para homens e 60% para mulheres. Esses resultados contradisseram nossa hipótese inicial de que o sexo feminino podia ser mais conservador que o sexo masculino no PGB e no PST: as mulheres produziram mais concordância verbal e, conseqüentemente, mais sujeitos nulos, embora todos estejam dentro da mesma faixa etária e exerçam a mesma função social: estudantes. Contudo, concluímos que o uso de sujeitos nulos e plenos não se trata da existência de uma variante de maior ou menor prestígio social, sendo assim, tanto homens quanto mulheres apresentaram um comportamento linguístico muito semelhante. Também atribuímos a diferença mínima no que concerne à ocorrência de sujeitos nulos, entre o sexo

masculino e o sexo feminino, na variedade guineense, ao fato de que, embora o crioulo guineense seja mais falado que o português, os falantes usaram mais nulos em decorrência do contexto, texto escrito, que exigiu maior grau de monitoramento do uso linguístico. Em relação a São Tomé e Príncipe, atribuímos ao resultado percentual acima o fato de que, em sendo o português não mais uma L2 para a maioria dos falantes são-tomenses, de acordo com Hagemeijer (2009), homens e mulheres são mais conservadores ao fazerem uso de estruturas mais próximas à norma europeia do português, confirmando-se assim nossa predição inicial de que os falantes são-tomenses fariam uso semelhante aos falantes do português europeu, no que concerne às ocorrências do sujeito nulo.

Quanto ao tempo de permanência no Brasil, o PGB mostrou resultados muito próximos em relação aos primeiros e últimos períodos, com 53% e 54% de ocorrências de sujeitos nulos, respectivamente. No que concerne ao PST, foram encontrados 62% e 64% de ocorrências de sujeitos nulos, em relação aos primeiros e últimos períodos, respectivamente. Esses resultados contradizem nossa hipótese inicial de que as restrições encontradas em PB para a distribuição de sujeitos nulos seriam evidenciadas na língua escrita dos estudantes dos últimos períodos de diferentes cursos da UNILAB em virtude do intenso contato do PGB e do PST com o PB

Atentamo-nos agora aos dados de fala, para considerar que nesta modalidade da língua pudemos observar mais significância nas variáveis, visto que os contextos de restrição de sujeitos nulos no PB (cf. DUARTE, 1995; CYRINO, KATO e DUARTE, 2000; DUARTE e KATO, 2014) foram observados com mais sensibilidade nos dados dessa modalidade. Quando comparados aos resultados de Duarte (1995) para o PE, temos que os dados de fala se distanciam da variedade europeia do português, visto que esta apresenta percentual de 66% para ocorrências de sujeitos nulos, enquanto encontramos um percentual de apenas 52% para sujeitos nulos.

Ao compararmos as variedades guineense e são-tomense do português individualmente, constatamos uma grande disparidade, dado que o PGB apresentou 61% de ocorrências de sujeitos plenos, o que é muito maior que o encontrado nos dados de escrita, enquanto o PST manteve uma aproximação entre os dados de fala e escrita. Em relação ao sexo dos informantes para o PGB, nossa hipótese de que as mulheres seriam mais conservadoras, mais uma vez, não se confirmou, pois, os homens fizeram uso mais expressivo das ocorrências de sujeitos nulos. Enquanto, para o PST as mulheres apresentaram resultados favoráveis ao uso de sujeitos nulos, com 65% do uso desses sujeitos, assim, corroborando

com o que foi visto, no PST, as mulheres tendem a fazer uso de estruturas mais próximas à norma europeia do português, sempre pensando que essa é a língua de ascensão nesse país (cf. HAGEMEIJER 2009).

O tempo de permanência no Brasil foi a variável mais significativa que encontramos, nos dados da modalidade de fala, visto que os alunos nos períodos finais produziram uma quantidade de sujeitos plenos muito significativa. No PGB, encontramos 71% de sujeitos foneticamente expressos, é um aumento de 20% ao resultado que havíamos encontrado para os dados de escrita. Esse resultado confirma nossa hipótese geral de uma interferência mais massiva do PB sobre o PGB, na medida em que a língua materna dos falantes do PGB também auxilia nessa interferência, pois o crioulo guineense é uma língua pro-drop. No tocante ao PST, constatamos mais um distanciamento do PB nessa variedade da língua, ao passo que os resultados encontrados nos dados de fala foram muito próximos aos encontrados nos dados de escrita e mantém resultados muito próximos entre os primeiros e últimos períodos.

Ademais, pontuamos as seguintes questões:

- i. A aquisição da língua portuguesa tem relação direta com nossa pesquisa, visto que os resultados entre PGB e PST divergem bastante entre si e isso se deve não apenas a interferência do PB, mas sim, à relação que os falantes têm com a variedade falada em seus países: em Guiné Bissau, o português é L2, a língua materna é o crioulo guineense e o português apenas é adquirido a partir da vida escolar; em São Tomé e Príncipe, o português é língua materna, aprendida em casa e na escola. Sendo assim, constatamos que o contato exercido entre o PGB e o crioulo guineense deixa registros naquela língua que também favorecem ao uso da posição preenchida de sujeito.
- ii. Encontramos alguns pontos de restrições ao uso do sujeito nulo na variedade guineense, que se aproximam dos contextos de restrições encontrados para o PB (CYRINO, DUARTE e KATO, 2000), (KATO e DUARTE, 2014), como em: a estrutura de duplicação de sujeito que ocorre nos dados de escrita e fala do PGB, mas que não foram encontrados no PST e que caracterizam uso de línguas não-pro-drop; uso de *você* genérico; uso da expressão pronominal *a gente* em concorrência com o pronome *nós*.
- iii. Vimos um enfraquecimento na morfologia de flexão verbal do PGB, visto que essa língua já produz ocorrências da expressão pronominal *a gente* no lugar semântico

da primeira pessoa do plural *nós*, faz uso também do *você* no lugar referente a segunda pessoa da gramatical, porém usa a desinência da terceira pessoa do singular. Enquanto que no PST, constatamos que a morfologia de flexão verbal é mais próxima ao que é visto no PE.

- iv. A assimetria entre língua falada e língua escrita é de extrema significância para nossa pesquisa, visto que os resultados apresentados entre os dados de ambas as modalidades foram divergentes. Houve uma grande ocorrência de produção de sujeitos nulos nos dados de escrita, enquanto que nos dados de fala houve preferência por sujeitos expressos. Ressaltamos que o grau de monitoramento linguístico utilizado para as duas modalidades também interfere nesses resultados, à medida que faz com o falante precise ter maior controle sobre os contextos de produção de sujeitos.

Por fim, os resultados condizem, na medida do possível com nossas hipóteses, contradizendo em alguns pontos e corroborando em outros. Percebemos divergências entre os resultados das variedades analisadas e, cada uma a seu modo, atestaram a interferência do PB sobre a produção de sujeitos nulos. Ao concluir a pesquisa que nos dispomos a fazer, deixamos como sugestão para pesquisas próximas, que novos dados sejam coletados, com uma outra proposta textual, a fim de que os alunos possam fazer uso mais expressivo de todas as pessoas gramaticais, com o intuito de produzir um quadro do paradigma flexional das variedades africanas que aqui foram estudadas.

REFERÊNCIAS

- ALEXANDRE, N.; DUARTE, I. e SANTOS, A. L. Infinitivos pessoais: uma viagem transatlântica. (2013). In D. Moura & M. Sibaldo (eds.). *Para a História do Português Brasileiro – Volume III: Estudos Comparativos entre o Português Brasileiro e Língua Crioulas de Base Lexical Portuguesa*, Tomo IV, 17-47, Maceió: EDUFAL.
- BARBOSA, P.; DUARTE, M. E. L. e KATO, M. A. A distribuição do sujeito nulo no português europeu e no português brasileiro. In: Correia, Clara Nunes & Anabela Gonçalves. *Actas do XVI Encontro Nacional da Associação Portuguesa de Linguística*. Lisboa: Colibri Artes Gráficas. 2001. p. 539-550.
- BARBOSA, P. A new look at the null subject parameter. In: COSTA, J. et al. (Eds.). *Proceedings of ConSOLE IV*. Leiden, nov. 1996. p. 375-395.
- _____. *Null Subjects*. Tese (doutorado em Linguística). MIT, 1995.
- _____.; DUARTE, M. E. L. e KATO, M. A. Null subjects in European and Brazilian Portuguese, *Journal of Portuguese Linguistics*, v. 4 (2). 2005.
- BAZENGA, A. 2016. *Sociedades insulares e identidade: aspetos linguísticos da “madeirensidade” no Atlântico lusófono*. In Grande Dicionário Enciclopédico da Madeira. Volume ante zero, ed. José Eduardo Franco, 216 - 225. ISBN: ISSN 1647-8088. Lisboa: CLEPUL/INCM, Suplemento da Revista Letras com Vida nº 7.
- BLEY-VROMAN, R. (1989) "What is the logical problem of foreign language learning? ". In: GASS, Suzan M. & SCHATER, Jacquelyn (eds) *Linguistic Perspectives on Second Language Acquisition*. Cambridge University Press.
- BRANDÃO, S. F. 2011. *Concordância nominal em duas variedades do português: convergências e divergências*. Veredas online, Atemática 1: 264-278.
- BRANDÃO, S. F. & VIEIRA, S. R. 2012. *A concordância nominal e verbal no Português do Brasil e no Português de São Tomé: uma abordagem sociolinguística*. Papiá 22: 7-33.
- BONVINI, E. Línguas africanas e português falado no Brasil. In: FIORIN, J. L.; PETTER, M. (org.) *África no Brasil*, 15-62. São Paulo: Contexto. 2008.
- BRAVIN DOS SANTOS, A. M. O português moçambicano entre o português brasileiro e o português europeu. *Filologia e Linguística Portuguesa*, 12(1), p. 27-39, 2010.
- CASTRO, P. P. *As construções interrogativas, de tópico e de foco na língua crioula de Guiné-Bissau*. 2013. Dissertação (Mestrado em Linguística) – Universidade Federal do Rio de Janeiro, Rio de Janeiro.
- CHOMSKY, N. *Syntactic Structures*. Mouton: The Hague, 1957.
- _____. *Aspectos da teoria da sintaxe*. 2. ed. Tradução de José António Meireles e Eduardo Paiva Raposo. Coimbra: Arménio Amado-Editor, 1978. (obra original (1965))

_____. *Lectures on government and binding*. Dordrecht : Foris Publications, 1981.

_____. *Some concepts and consequences of the theory of government and binding*. v. 6, MIT Press, 1982.

_____. *Knowledge of language: its nature, origin and use*. London : Praeger Publishers, 1986.

_____. *O conhecimento da língua: sua natureza, origem e uso*. Lisboa: Caminhos. 1994.

CLAHSEN, H.; MUYSKEN, P. The availability of universal grammar to adult and child learners: A study of the acquisition of German Word order. *Second Language Research*. n. 2, p. 93-119, 1986.

CLAHSEN, H. Parameterized grammatical theory and language acquisition: a study of the acquisition of verb placement and inflection by children and adults. In: S. FLYNN; W. O'NEIL. (Eds.). *Linguistic theory in second language acquisition*. Dordrecht: Kluwer, 1988. p. 47-75.

COUTO, H. H. (1989) *O Crioulo guineense em relação ao português e às línguas nativas*. *Linguística XXIX*. p. 107-128.

COUTO, H. H. e EMBALÓ, F. *Literatura, língua e cultura na Guiné-Bissau: um país da CPLP*. In: *Papia*. n.20, 2010

CRUZ, A. V. *Abordagem Comunicativa – Enfoque na Competência Oral na Língua Segunda. Caso da Guiné-Bissau*. 2013. Dissertação (Mestrado em Português Segunda Língua) – Universidade do Porto, Porto.

CYRINO, S. *O objeto nulo no português do Brasil*. Londrina: Ed. Da UEL, 1997.

CYRINO, S.; DUARTE, M. E. & KATO, M. A. “Visible subjects and invisible clitics in Brazilian Portuguese”. In: KATO, M. A. & NEGRÃO, E.V. (eds.). *Brazilian Portuguese and the Null Subject Parameter*. Frankfurt-Madrid: Vervuert-Iberoamericana, pp. 55-104. 2000.

DUARTE, M. E. L. Do pronome nulo ao pronome pleno: a trajetória do sujeito no português do Brasil. In: ROBERTS, I & M. A. KATO (Org.). *Português Brasileiro: uma viagem diacrônica*, 107-128. Campinas: Editora da unicamp. 1993.

_____. *A Perda do Princípio “evite pronome” no português brasileiro*. Tese – (Doutorado em Linguística), unicamp, Campinas. 1995.

_____. The loss of the Avoid Pronoun principle in Brazilian Portuguese. In: Kato, M. A.; Negrão, E. V. (Eds). *Brazilian and the Null Subject Parameter*. Frankfurt-Madrid: Vervuert-Iberoamericana, 2000, p. 17-36.

DUARTE, D. A. Apud, Herculano Simplício Rodrigues. *Cabo Verde – o Português e o Crioulo em presença: proposta de uma abordagem metodológica*, Universidade de Beira Interior, Covilhã. 2006.

EUBANK, L. On the transfer of parametric values in L2 development. *Language Acquisition*. n. 3, p. 182-208, 1993.

EUBANK, L. Optionality and the initial state in L2 development. In: T. HOEKSTRA; B. SCHWARTZ (Eds.). *Language acquisition studies in generative grammar*. 1994. p. 369-88.

_____. Negation in early German-English Interlanguage: more 'valueless features' in the L2 initial state. *Second Language Research*. n. 12, p. 73-106, 1996.

FODOR, J. *The modularity of mind*. Cambridge, Mass.: The MIT Press, 1983.

FONSECA, S. P. B. *Aquisição e aprendizagem da referência nominal no contexto do português língua não materna na Guiné-Bissau*. 2012. Dissertação (Mestrado em Português Língua Não-Materna) – Universidade Aberta, Lisboa.

FLYNN, S. (1987). *A parameter setting model of L2 acquisition*. Dordrecht: Reidel Publishing.

_____. Government-binding: Parameter setting in second language acquisition. In: *Crosscurrents in second language acquisition and linguistic theories*. C. FERGUSON; T. HUEBNER (Eds.). Amsterdam: John Benjamins, 1991.

GALVES, C. M. C. "A sintaxe do português brasileiro". In: *Cadernos de Lingüística e Teoria da Literatura – Ensaios de Lingüística* 7, 1987.

_____. *Ensaios sobre as gramáticas do português*. Campinas: Editora da UNICAMP. 2001

_____. O enfraquecimento da concordância no português brasileiro. In: ROBERTS, I; GONÇALVES, P. 1990. *A construção de uma gramática de português em Moçambique: aspectos da estrutura argumental dos verbos*. Dissertação de Doutoramento. Lisboa: Universidade de Lisboa.

GONÇALVES, P. 2005. Falsos Sucessos no Processamento do Input na Aquisição de L2: Papel na Ambiguidade na Génese no Português de Moçambique. *Revista da ABRALIN* 4 (1) (2), 47-73.

GONÇALVES, R. M. G. *Propriedades de subcategorização verbal no português de São Tomé*. 2010. Dissertação de Mestrado, universidade de Lisboa, Lisboa.

HAGEMEIJER, T. 2007. *Clause structure in Santome*. Tese de Doutoramento. Lisboa: Universidade de Lisboa.

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

HOLMBERG, A. Is there a little pro? Evidence from Finnish. *Linguistic Inquiry*, v. 36, p. 533-564, 2005.

HOLMBERG, A.; NAYUDU, A.; SHEEHAN, M. *Three Partial null-subject languages: a comparison of Brazilian Portuguese, Finnish and Marathi*. *Studia Lingüística*, n. 63, p. 59-97, 2009.

HUANG, J. C. T. (1984). *On the distribution and reference of the empty categories*. *Linguistic Inquiry*, 15. 531 -574.

_____. Pro-drop in Chinese: a generalized control theory. In: JAEGLI, O.; SAFIR, K. J. *The null subject parameter*. Dordrecht, London: Kluwer Academic Publishers. p. 185-214, 1989.

JAEGLI, O. & SAFIR, K. J. (1989). The Null Subject Parameter and Parametric Theory. In: O. Jaeggli & K. J. Safir (eds.) *The Null Subject Parameter*. Dordrecht: Kluwer. 1-44.

KATO, M. A. (org.). *Português brasileiro: uma viagem diacrônica*, p. 387- 408. Campinas: Editora da unicamp. 1993.

_____. *Português brasileiro falado: aquisição em contexto de mudança linguística*. Congresso Internacional sobre o Português. Lisboa. Digitado. 1994.

KATO, M. A.; DUARTE, M. E. L. A variação entre construções finitas pessoais e impessoais no português brasileiro. *Web-Revista SICIODIALETO*. v. 4, 12: 153-177. 2014a.

KATO, M. A.; DUARTE, M. E. L. 2014b. Restrições na distribuição de sujeitos nulos no português brasileiro. *Revista Veredas*. v. 18, 1, p. 1-22. 2014b.

KATO, M. A.; NEGRÃO, E. V. (eds.). *Brazilian Portuguese and the null subject parameter*. Madrid: Vervuert, Iberoamericana, 2000.

KATO, M. Child L2 acquisition: an insider account. In: N. MÜLLER (Ed.). *Vulnerable domains in Multilingualism*. John Benjamins, 2003. p. 93-271.

LAKATOS, E. M.; MARCONI, E. M. A. *Fundamentos de metodologia científica* 1. - 5. ed. - São Paulo : Atlas 2003.

LEFFA, V. J. Metodologia do ensino de línguas. In BOHN, H. I.; VANDRESEN, P. *Tópicos em lingüística aplicada: O ensino de línguas estrangeiras*. Florianópolis: Ed. da UFSC, p. 211-236, 1988.

LENNEBERG, E. *Biological Foundations of Language*. New York: Wiley, 1967.

LOPES, F. J.; CAMPOS, E. A. A expressão pronominal no português brasileiro e no português falado em Cabo Verde — trilhando possíveis (as) simetrias. *PAPIA*, São Paulo, 25(2), p. 319-345, 2015.

LORENZINO, G. A. 1996. Uma avaliação sociolinguística sobre São Tomé e Príncipe. In: DUARTE, I. e LEIRIA, I. (orgs), *Actas do Congresso Internacional sobre o Português*, 435-449. Vol. II. Lisboa: APL e Edições Colibri.

LUCCHESI, D. (2006) *Parâmetros sociolinguísticos do português brasileiro*. Revista da ABRALIN, v. 5, n. 1, p. 83-112.

LUCCHESI, D.; BAXTER, A.; RIBEIRO, I. (2009) *O português afro-brasileiro*. SciELO-Edufba.

LUNGUINHO, M. V.; RESENES, M. S.; NEGRÃO, E. V. Pesquisa em sintaxe gerativa: pressupostos teóricos, procedimentos metodológicos e questões. In: GONÇALVES, Adair Vieira; GÓIS, Marcos Lúcio de Sousa. (org) *Ciências da Linguagem: o fazer científico?*, 119-161. São Paulo: Mercado das Letras, 2012.

LYONS, J. *Lingua(gem) e lingüística: uma introdução*. Tradução de Marilda Winkler Averburg e Clarisse Sieckenius de Souza. Rio de Janeiro: LTC, 1987.

MAGALHÃES, T.; SANTOS, A. L. *As respostas verbais e a frequência de sujeito nulo na aquisição do português europeu e do português brasileiro*. Texto apresentado no primeiro CLIP, UFSC, 2004.

MAGALHÃES, T. M. V. *O sistema pronominal sujeito e objeto na aquisição do português europeu e do português brasileiro*. 2006. Tese (doutorado em Linguística) – Instituto de Estudos da Linguagem da Universidade Estadual de Campinas, Campinas.

MARQUES, M. E. *Português, Língua Segunda*, Universidade Aberta, Lisboa. 2003.

MATTOS E SILVA, R. V. “Diversidade e Unidade: A Aventura Linguística do Português”. Revista ICALP, vol. 11, 1988, 60-72; Revista ICALP, vol. 12/13, 1988, 13-28.

MIOTO, C.; FIGUEIREDO SILVA, M. C.; LOPES, R. E. V. *Novo Manual de Sintaxe*. 3. Ed. Florianópolis: Insular, 2013.

MOTA, M. A. *Para uma tipologia da concordância sujeito-verbo, em português falado: contributos do português de Luanda e de Cabo Verde*. Cuadernos de La ALFAL, n. 7, p. 17-35, mar. 2015.

MOTA, M. A.; MIGUEL, M.; MENDES, A. A concordância de p6 em português falado. Os traços pronominais e os traços de concordância. *Papia*, v. 22, n. 1, p. 161-187, 2012.

OLIVEIRA, M. S. D.; SANTOS, E. F. Pronomes nulos na posição de sujeito no português de Angola – um estudo preliminar. *Filologia e Linguística Portuguesa*, n. 9, p. 85-101, 2007.

OLIVEIRA, V. M. S. *A expressão do sujeito no português de Moçambique*. 2016. Dissertação (Mestrado em Linguística) – Universidade Estadual de Campinas, Campinas.

OLIVEIRA, G. F.; SILVA, C. R. T. A distribuição do nulo: uma análise comparative entre o português europeu e o português de São Tomé. In: SILVA, J. E. N.; ARRUDA, L. S.; ALVES, D. M. V. (orgs) *Línguas e Culturas: contatos, conflitos, nomadismos* [livro eletrônico] Faculdade de Letras, UFRJ, 2018

- OMENA, N. P. A referência á primeira pessoa do discurso no plural. In: SILVA, G. M. de O. e; SCHERRE, M. M. P. *Padrões Sociolingüísticos: análise de fenômenos variáveis do português falado na cidade do Rio de Janeiro*. Rio de Janeiro: Tempo Brasileiro: Departamento de Linguística e Filologia, UFRJ, 1996, p. 183-215.
- PEREIRA, D. *Crioulos de base portuguesa*. Lisboa: Caminho, 2007.
- PETTER, M. M. T. Aspectos morfossintáticos comuns ao português angolano, brasileiro e moçambicano. *Papia*, n. 19, p. 201-220, 2009.
- PETTER, M. *Introdução à linguística africana*. São Paulo: Contexto, 2015.
- RAPOSO, E. P. (1992). *Teoria da gramática: a faculdade da linguagem*. Lisboa: Editorial Caminho.
- _____. (1993). *Categorias funcionais na gramática gerativa*. DELTA, vol. 9, nº 2. 217-336.
- RIZZI, L. *Issues in Italian Syntax*. Dordrecht: Foris, 1982.
- RIZZI, L. Null subjects in Italian and the theory of pro. *Linguistic Inquiry*, v. 17, n. 3, p. 501-558, 1986.
- ROBERTS, I. (1993) "O português brasileiro no contexto das línguas românicas". In: ROBERTS, I & KATO, M. A. (orgs) *Português Brasileiro: urna viagem diacrônica*. Campinas: Ed. da UNICAMP. 409-421.
- ROBERTS, I. *Creoles, markedness and the Language Bioprogram Hypothesis*. Estudos Linguísticos e Literários, Salvador, n.19, p.11-24, 1997.
- ROBERTS, I.; HOLMBERG, A. Introduction: parameters in minimalist theory. In: BIBERAUER, T.; HOLMBERG, A.; ROBERTS, I.; SHEEHAN, M. *Parametric Variation: Null Subjects in Minimalist Theory*. Cambridge: Cambridge University Press, 2010. p. 1-57.
- RODRIGUES, C. (2004). *Impoverished morphology and A-movement out of case domains*. Tese de Doutorado. Maryland: University of Maryland.
- SANI, F. *O Lugar e as relações do português como língua oficial na escola da Guiné-Bissau*. 1996. Dissertação (Mestrado em Linguística) – Universidade Estadual de Campinas, Campinas.
- SCHERRE, M. M. P.; NARO, A. J. *Mudança sem mudança: a concordância de número no português brasileiro*. SCRIPTA, v. 9, n 18, p. 107-129, 2006.
- SCHWARTZ, B. and SPROUSE, P. Word order and nominative case in nonnative language acquisition: A longitudinal study of (L1 Turkish) German inter-language. In: T. HOEKSTRA; B. SCHWARTZ (Eds.). *Language acquisition studies in generative grammar*. Amsterdam: John Benjamins, 1994. p. 68-317.

SCHWARTZ, B. and SPROUSE, P. (1996). "L2 cognitive states and Full Transfer/ Full Access model". *Second Language Research* 12 (1): 40-72.

SILVA, C. R. T. *A natureza de AGR e suas implicações na ordem VS: um estudo comparativo entre o português brasileiro e o português europeu*. 2004. Tese (Doutorado em Linguística) – Universidade Federal de Alagoas, Maceió.

SILVA, C. R. T.; CARVALHO, D. S.; ZIOBER, F. M. Licenciamento de duplos sujeitos em variedades do Português: pessoa, definitude e estrutura de traços. *Letras escreve*. Macapá, v. 7, n. 2, 2º semestre, 2017.

SNOW, C. Relevance of the notion of a critical period to language acquisition. In: BORNSTEIN, M. H. (Ed.). *Sensitive Periods in Development: Interdisciplinary Perspectives*. Hillsdale N. J.: Lawrence Erlbaum, 1987. p. 183-210.

STROZER, J. *Language acquisition after puberty*. Washington, DC: Georgetown University Press, 1992.

TARALLO, F. & KATO, M. A. "Harmonia trans-sistêmica: variação inter- e intralingüística". *Preedição 5*. Campinas, Unicamp, pp. 315-353. 1989.

TARALLO, F.. "Diagnosticando uma gramática brasileira: o português d'aquém e d'além mar ao final do século XIX". In: ROBERTS, I. & KATO, M. A. (orgs.). *Português brasileiro: uma viagem diacrônica*. Campinas, SP: Pontes, 1993, pp. 69-105.

TEIXEIRA, E. P. O pronome você no português de Luanda. In: LIMA- Hernandez, Maria Célia et al. *A língua portuguesa no mundo*. São Paulo: FFLCH-USP, 2008.

TEIXEIRA, E. P. (2012) *A representação do sujeito pronominal no português popular angolano*. PAPIA-Revista Brasileira de Estudos Crioulos e Similares, v. 22, n. 1, p. 141-159

XAVIER, G. R. *Português brasileiro como segunda língua: um estudo sobre o sujeito nulo*. 2006. Tese (Doutorado em Linguística) – Universidade Estadual de Campinas, Campinas.

VAINIKKA, A. and YOUNG-SCHOLTEN, M. (1996): "Gradual development of L2 phrase structure ". In: *Second Language Research*. 12 (1): 7-39.

VIANNA, J.; LOPES, C.. A variação entre nós e a gente: uma comparação entre o português europeu e o brasileiro. *Revista do GELNE*, v. 14, p. 95-116, 2011.

ZILLES, A. M. S. O que a fala e a escrita nos dizem sobre a avaliação social do uso de a gente? In: *Letras de Hoje*. Porto Alegre, v.42, n.2, p. 27-44, junho de 2007.