UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE ARTES E COMUNICAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM LETRAS

ROGÉRIO DE OLIVEIRA JÚNIOR

CONCORDÂNCIA E PARÂMETRO DO SUJEITO NULO EM REDAÇÕES PRODUZIDAS POR TIMORENSES

Recife
2021

ROGÉRIO DE OLIVEIRA JÚNIOR

CONCORDÂNCIA E PARÂMETRO DO SUJEITO NULO EM REDAÇÕES PRODUZIDAS POR TIMORENSES

Dissertação apresentada ao Programa de Pós Graduação em Letras da Universidade Federal de Pernambuco, como requisito parcial à obtenção do Título de Mestre em Letras.

Área de concentração: Linguística

Orientadora: Profª. Drª. Claudia Roberta Tavares Silva

Recife

2021

Catalogação na fonte
Bibliotecária Jéssica Pereira de Oliveira – CRB-4/2223

O48c Oliveira Júnior, Rogério de
Concordância e Parâmetro do Sujeito Nulo em redações produzidas por timorenses / Rogério de Oliveira Júnior. – Recife, 2021.
137p.: il.

Orientadora: Claudia Roberta Tavares Silva.
Dissertação (Mestrado) – Universidade Federal de Pernambuco. Centro de Artes e Comunicação. Programa de Pós-Graduação em Letras, 2021.

Inclui referências.

1. Parâmetro do Sujeito Nulo. 2. Português de Timor-Leste. 3. Multilinguismo. 4. Contato Linguístico. I. Silva, Claudia Roberta Tavares (Orientadora). II. Título.

410 CDD (22. ed.) UFPE (CAC 2021-168)

ROGÉRIO DE OLIVEIRA JÚNIOR

CONCORDÂNCIA E PARÂMETRO DO SUJEITO NULO EM REDAÇÕES PRODUZIDAS POR TIMORENSES

Dissertação apresentada ao Programa de Pós Graduação em Letras da Universidade Federal de Pernambuco, como requisito parcial à obtenção do Título de Mestre em Letras.

Aprovada em: 31/05/2021.

BANCA EXAMINADORA

_____

Profª. Drª. Claudia Roberta Tavares Silva (Orientadora)

Universidade Federal de Pernambuco

_____

Profª. Drª. Telma Moreira Vianna Magalhães (Examinadora Externa)

Universidade Federal de Alagoas

_____

Profª. Drª. Silvia Rodrigues Vieira (Examinadora Externa)

Universidade Federal do Rio de Janeiro

AGRADECIMENTOS

Agradeço primeiramente a Jeová Deus por ter me dado a vida e a oportunidade de estudar esse objeto complexo e maravilhoso que ele criou, a linguagem. Por todo presente perfeito, que desce do Pai das luzes celestiais, sou muito grato.

Agradeço à professora Cláudia Roberta, minha querida orientadora, pelo lindo trabalho que realiza pesquisando a linguagem e pela paciência, cuidado e generosidade que teve comigo nesses anos desde a graduação.

Agradeço a Fernanda Miguel, grande incentivadora, pesquisadora de mão cheia, que bem lembrou da importância de ter disciplina e diligência ao lidar com tantas tarefas.

Agradeço a minha família, especialmente a minha mãe, Andréa Carla, por me ajudar a ver desde cedo que a educação é algo muito importante e o que ajuda a mudar a realidade, a minha e a do outro.

Agradeço ao querido amigo Felipe Siongo pelo empenho na ajuda com o mapeamento dos dados no *corpus* da pesquisa.

Agradeço ao professor André Pedro pelas importantes ajudas no tratamento quantitativo dos dados a partir da utilização do programa GoldVarb X.

Agradeço à professora Cristiane Abreu pelos seus ensinamentos e pela paixão pela linguagem e pelas pessoas a quem ensinamos.

Agradeço à professora Tatiana Luna pela orientação desde o ensino médio, por me iniciar na pesquisa e pelo incentivo sem igual pelas Letras.

Agradeço à colega Maria Joseane pela generosidade em compartilhar comigo as redações coletadas na UNILAB para que eu pudesse constituir o *corpus* desta pesquisa.

Agradeço à UNILAB pela oportunidade de realizar esta pesquisa com os estudantes timorenses que passaram por ali, pela importância que é essa instituição pública brasileira na formação de pessoas, em especial, de estrangeiros falantes de português.

Agradeço ao povo de Timor-Leste por tudo o que passou e tem passado para firmar o seu caminhar nesse mundo, nem sempre fácil, pelo empenho destacado em valorizar o português como língua sua e pela riqueza linguística que possui.

Agradeço à UFRPE e aos seus colaboradores pelo início da minha caminhada acadêmica. Pela riqueza, carinho e generosidade das pessoas dessa instituição, a “Ruralinda” sempre será lembrada.

Agradeço ao professor Alberto Poza pelo apoio durante o mestrado e a consideração especial para que eu pudesse integralizar as disciplinas.

Agradeço a Claudyvanne Silva e a Jozaías Santos, servidores do PPGL, pelo bom tratamento e o pronto atendimento das minhas demandas durante o curso.

Agradeço aos professores Antônio Xavier, Cláudia Roberta, Vicente Masip, Alberto Poza e Marcelo Sibaldo por ministrarem as disciplinas que cursei no Programa, que foram indispensáveis para a minha formação enquanto linguista e pesquisador.

Agradeço às professoras Telma Magalhães e Silvia Vieira por fazerem parte da minha banca examinadora e pelas importantes contribuições à pesquisa.

Agradeço aos colegas Yasmin Galindo, Danillo Melo e Edite Consuelo pelos importantes incentivos e pela companhia na caminhada acadêmica.

Os erros são todos meus.

“Jeová lhe disse: Quem designou a boca ao homem? Não sou eu, Jeová?” (BÍBLIA, Êxodo, 4, 11)

RESUMO

Na presente dissertação, realizo um estudo sobre o Português de Timor-Leste (PTL), analisando em textos escritos como se comporta o Parâmetro do Sujeito Nulo (PSN), seguindo a perspectiva do Modelo de Princípios e Parâmetros (CHOMSKY, 1981 e seguintes). Com base nesse parâmetro, investigarei como a posição de sujeito está sendo preenchida nessa variedade não-europeia do português, sendo a questão central: há mais sujeitos não-realizados foneticamente (sujeitos nulos), indo na direção do português europeu, ou mais sujeitos realizados foneticamente (sujeitos plenos)? (CHOMSKY, 1981; RIZZI, 1986, DUARTE, 1995). Mais recentemente, pesquisadores gerativistas, ao analisarem diferentes línguas, concluem que o português europeu (PE) é uma língua de sujeito nulo consistente, ao passo que o português brasileiro (PB), uma língua de sujeito nulo parcial (HOMLBERG; NAYUDU; SHEEHAN, 2009; KATO; DUARTE, 2014). Vale referir que, no Timor-Leste, o português atualmente não é língua materna da população e coexiste com outras línguas locais, sendo o tétum, uma língua de sujeito não-nulo, a mais usada (ALBUQUERQUE, 2011). O *corpus* é composto de 582 dados, extraídos de redações produzidas por timorenses em um exame vestibular para ingresso na Universidade da Integração Internacional da Lusofonia Afro-brasileira (UNILAB), localizada na cidade de Redenção-CE. Esses dados foram submetidos a tratamento quantitativo após serem codificados, consoante às variáveis selecionadas (dependente e independentes). Os resultados revelam uma produção de sujeitos plenos e nulos muito próxima, o que, consoante às demais características morfossintáticas analisadas, possibilita-me considerar que o PTL ainda não apresenta as propriedades de uma língua de sujeito nulo como o PE, uma conclusão que decorre provavelmente da influência do contexto multilíngue em que PTL está inserido, conforme já verificado em outras pesquisas (SANTOS, 2009; DEUS, 2009; MARTINS, 2016; GUIMARÃES, 2017, ALBUQUERQUE; RAMOS, 2020).

Palavras-chave: Parâmetro do Sujeito Nulo; Português de Timor-Leste; Multilinguismo; Contato Linguístico.

ABSTRACT

In this dissertation, I conduct a study on the Portuguese of Timor-Leste (PTL), analyzing in written texts how the Null Subject Parameter (PSN) behaves, following the perspective of the Principles and Parameters Model (CHOMSKY, 1981 and following). Based on this parameter, I will investigate how the subject position is being filled in this non-European variety of Portuguese, and the central question is: there are more non-realized subjects phonetically (null subjects), going in the direction of European Portuguese, or phonetically performed subjects (full subjects)? (CHOMSKY, 1981; RIZZI, 1986, DUARTE, 1995). More recently, generative researchers, when analyzing different languages, concluded that European Portuguese (PE) is a consistent null subject language, while Brazilian Portuguese (PB), a partial null subject language (HOMLBERG; NAYUDU; SHEEHAN, 2009; KATO; DUARTE, 2014). It is worth mentioning that, in Timor-Leste, Portuguese is not currently the mother tongue of the population and coexists with other local languages, being Tetum, a non-null subject language, the most used (ALBUQUERQUE, 2011). The corpus is composed of 582 data, extracted from essays produced by Timorese in an exam for admission to the Universidade da Integração Internacional da Lusofonia Afro-Brasileira (UNILAB), located in the city of Redenção-CE (Brazil). These data were submitted to quantitative treatment after being coded, according to the selected variables (dependent and independent). The results reveal a remarkably close production of full and null subjects, which, according to the other morphosyntactic characteristics analyzed, allows me to consider that the PTL still does not present the properties of a null subject language such as PE. This conclusion probably follows the influence of the multilingual context in which PTL is inserted, as verified in other research (SANTOS, 2009; DEUS, 2009; MARTINS, 2016; GUIMARÃES, 2017, ALBUQUERQUE; RAMOS, 2020).

Keywords: Null Subject Parameter; East Timor Portuguese; Multilingualism; Linguistic Contact.

LISTA DE ILUSTRAÇÕES

Figura 1 - Níveis de representação, de acordo com Princípios e Parâmetros.....50
Figura 2 - Processo de aquisição no período crítico.....69
Gráfico 1 - Realização do sujeito pleno.....98
Gráfico 2 - Distribuição dos sujeitos no PTL.....105
Gráfico 3 - Tipo de sujeito por sexo do escrevente.....111

LISTA DE QUADROS

Quadro 1 - Línguas crioulas de base portuguesa na África e na Ásia.....18
Quadro 2 - Paradigma de flexão verbal do PE e do Tétum.....24
Quadro 3 - Paradigma de flexão verbal do PB e do PE.....31
Quadro 4 - Distribuição dos dados com e sem marca verbal de 3ª pessoa plural no PM.....39
Quadro 5 - Número de ocorrências das formas pronominais na posição de sujeito no PA.....41
Quadro 6 - Quantitativo geral de sujeitos plenos e nulos nos corpora do PA.....43
Quadro 7 - Quantitativo geral de sujeitos plenos e nulos nos corpora do PST.....43
Quadro 8 - Quantitativo geral de sujeitos plenos e nulos nos corpora do PGB.....44
Quadro 9 - Desenvolvimento de L1 e L2.....72
Quadro 10 - Constituição do *corpus*.....81
Quadro 11 - Fatores controlados das variáveis.....90
Quadro 12 - Totais por tipo de verbo nos dados gerais.....94
Quadro 13 - Totais por tipo de oração nos dados gerais.....95
Quadro 14 - Totais por tipo de referência semântica do sujeito nos dados gerais.....96
Quadro 15 - Concordância verbal do sujeito pleno.....102
Quadro 16 - Concordância verbal do sujeito pleno pronominal.....106
Quadro 17 - Distribuição de sujeitos nulos e plenos por tipo de verbo.....107
Quadro 18 - Distribuição de sujeitos nulos e plenos por tipo de oração.....109
Quadro 19 - Distribuição de sujeitos nulos e plenos por referência semântica do sujeito.....110
Quadro 20 - Distribuição de sujeitos nulos e plenos de acordo com o sexo do escrevente.....111
Quadro 21 - Resultados da realização do sujeito em Santos (2009).....114
Quadro 22 - Resultados da concordância verbal em Santos (2009).....115
Quadro 23 - Tipo de verbo das orações no infinitivo.....119
Quadro 24 - Referência semântica do sujeito nas orações no infinitivo.....120

LISTA DE ABREVIATURAS E SIGLAS

ACC	Caso Acusativo
AGR	Concordância
CPLP	Comunidade dos Países de Língua Portuguesa
CP	*Complementizer Phrase*
DP	*Determiner Phrase*
EPP	*Extended Projection Principle*
GU	Gramática Universal
FL	Faculdade da Linguagem
HAN	Hipótese de Acesso Nulo
HAP	Hipótese de Acesso Parcial
HAT	Hipótese de Acesso Total
IP	*Inflectional Phrase*
LE	Língua estrangeira
LM	Língua Materna
LNM	Língua Não Materna
LO	Língua oficial
NOM	Caso Nominativo
NS	*Null Subject*
NURC	Projeto da Norma Urbana Linguística Culta
OBL	Caso Oblíquo
PA	Português de Angola
PB	Português brasileiro
PE	Português europeu
PM	Português de Moçambique
PGB	Português de Guiné-Bissau
PSN	Parâmetro do Sujeito Nulo
PST	Português de São Tomé
PTL	Português de Timor-Leste
SN	Sintagma Nominal
UNILAB	Universidade da Integração Internacional da Lusofonia Afro-Brasileira
VP	*Verbal Phrase*

SUMÁRIO

1 PONTO DE PARTIDA.....14
1.1 INTRODUÇÃO.....14
1.2 CONTATO LINGUÍSTICO DO PORTUGUÊS COM OUTRAS LÍNGUAS.....16
1.2.1 A realidade linguística de Timor-Leste.....20
1.3 CARACTERIZAÇÃO DO OBJETO DE ESTUDO.....28
1.3.1 Incursões no Parâmetro do Sujeito Nulo.....28
1.3.2 O parâmetro do sujeito nulo em variedades do português.....29
1.4 OBJETIVOS.....45
1.4.1 Geral.....45
1.4.2 Específicos.....45
1.5 HIPÓTESES.....45
1.5.1 Geral.....45
1.5.2 Específicas.....46
1.6 CONCLUSÃO DO CAPÍTULO.....46
2 BASE TEÓRICA DA PESQUISA.....47
2.1 A TEORIA GERATIVA: PRESSUPOSTOS BÁSICOS.....47
2.1.1 A hipótese inatista.....47
2.1.2 A Faculdade da Linguagem.....52
2.2 O PARÂMETRO DO SUJEITO NULO.....56
2.3 A AQUISIÇÃO DA LINGUAGEM.....67
2.3.1 Aquisição *versus* aprendizagem.....70
2.3.1.1 O caso específico de Timor-Leste.....75
3 PROCEDIMENTOS METODOLÓGICOS.....77
3.1 TIPO DE MÉTODO.....77
3.2 COLETA DE DADOS.....78
3.3 PARTICIPANTES DA PESQUISA.....80
3.4 CONSTITUIÇÃO DO *CORPUS*.....80
3.5 SELEÇÃO DAS VARIÁVEIS.....81
3.5.1 Variáveis.....82
3.5.1.1 Variável dependente.....82
3.5.1.2 Variáveis linguísticas.....82

3.5.1.3 Variável extralinguística.....89
3.6 TIPO DE ANÁLISE.....90
4 COMPREENDENDO A CONCORDÂNCIA E O SUJEITO NULO NOS DADOS ESCRITOS POR ESTUDANTES TIMORENSES.....92
4.1 RESULTADOS GERAIS.....93
4.1.1 Resultados por variável.....93
4.1.1.1 Tipo de verbo.....93
4.1.1.2 Tipo de oração.....95
4.1.1.3 Pessoa e número do sujeito.....96
4.1.1.4 Realização do sujeito pleno.....97
4.1.1.5 Posição do sujeito pleno.....100
4.1.1.6 Duplicação do sujeito pleno.....101
4.1.1.7 Morfologia de flexão verbal.....102
4.2 RESULTADOS GERAIS PARA O SUJEITO NULO.....104
4.2.1 Resultados por variável.....104
4.2.1.1 Tipo de sujeito.....104
4.2.1.2 Morfologia de flexão verbal.....106
4.2.1.3 Tipo de verbo.....107
4.2.1.4 Tipo de oração.....108
4.2.1.5 Pessoa e número do sujeito.....110
4.2.1.6 Sexo do escrevente.....111
4.3 DISCUSSÃO DOS RESULTADOS.....112
4.4 O CASO DOS SUJEITOS EM ORAÇÕES INFINITIVAS: BREVES INCURSÕES.....118
4.4.1 Ocorrência e não ocorrência de infinitivo nos textos do *corpus*.....122
5 CONSIDERAÇÕES FINAIS.....127
REFERÊNCIAS.....129

1 PONTO DE PARTIDA

1.1 INTRODUÇÃO

Os empenhos feitos com o fim de estudar a primorosa engenhosidade da comunicação humana têm-nos levado a alcançar novos patamares na compreensão não apenas dos processos comunicativos, mas também da própria constituição do Homem como ser social. A Linguística Moderna, por exemplo, estabeleceu programas de investigação promissores, que têm percebido a língua a partir de suas várias facetas e realizações. Algumas dessas facetas, por um lado, observam o empreendimento linguístico de fora (a chamada "Linguística Externa"). Elas, de maneira geral, seguindo uma perspectiva funcional, compreendem a língua como fruto de atividades socialmente relevantes, sendo constituída e modificada no uso linguístico dos falantes que atuam e interagem ativamente no processo comunicativo. Nesse sentido, a multidisciplinaridade da Linguística com as Ciências Sociais tem sido muito profícua, sobretudo, em compreender o Homem como um ser eminentemente social que se constitui na e pela língua.

Por outro lado, esforços notáveis têm sido feitos desde o século XX para investigar a língua de dentro (a chamada "Linguística Interna"), que, portanto, é entendida sob uma perspectiva formal. Nesse domínio, destaca-se o viés teórico de que inatamente todo o Ser Humano tem a predisposição genética para adquirir uma língua natural (hipótese inatista). Assim, os processos biológicos, mentais garantem a proeza não só de usar a língua, mas também de adquiri-la e expressá-la. Digno de nota é o esforço encabeçado, desde o início da segunda metade do século passado, pelo linguista Noam Chomsky, a respeito da constituição e do funcionamento surpreendente da Faculdade da Linguagem, considerado um órgão da mente-cérebro responsável por propiciar a aquisição de uma língua natural através de um Dispositivo de Aquisição da Linguagem, a saber: a Gramática Universal (GU). Esse dispositivo inato traz em si os princípios, ou regularidades universais que são comuns a todas as línguas, e também parâmetros que permitem a diferenciação entre línguas particulares. Assim, a surpreendente semelhança na gramática das línguas é fruto desse aparato natural, que permitirá às crianças adquirir uma língua específica, por

fixar os valores paramétricos dela. Desde 1981, o preenchimento da posição sujeito vem sendo analisada por diversos pesquisadores, permitindo uma compreensão das propriedades do **Parâmetro do Sujeito Nulo** (doravante, PSN) (CHOMSKY, 1981, 1986; RIZZI, 1982, 1985). Num entendimento inicial, foi defendida a correlação entre morfologia de flexão verbal (*agreement* (AGR) “concordância”) rica e a produção de sujeitos nulos, a exemplo de línguas como o italiano, o catalão, o português europeu (PE). (TARALDSEN, 1978; RIZZI, 1986, DUARTE, 1995).

Diante do exposto, o trabalho aqui empreendido por mim se destina a mapear a distribuição de sujeitos nulos e plenos em um *corpus* escrito do português leste-timorense (PTL), visando verificar se a natureza de *agreement* interfere na produção de sujeitos nulos e plenos nos dados submetidos à análise. Para tanto, os dados que compõem esse *corpus* foram extraídos de redações produzidas por timorenses já escolarizados que se submeteram a um exame vestibular para ingresso na Universidade de Integração Internacional da Lusofonia Afro-Brasileira (UNILAB) localizada na cidade de Redenção-CE, possibilitando observar como é a sua gramática estável.

Ademais, vale referir que o contexto linguístico em que está inserido o PTL é muito rico e variado, reunindo, em um pequeno país, várias línguas diferentes. Esse contato linguístico pode ter feito com que o português ali utilizado tenha características também dessas línguas.

Na investigação, visando a um melhor entendimento de como o PSN comporta-se nos dados do PTL em análise, empreguei-me a ampliar a discussão ao revisitar resultados de pesquisas sobre esse parâmetro e a natureza da concordância em diferentes variedades do português. Essa comparação faz-se necessária sempre que possível ao longo desta dissertação, para que se possa compreender se o PTL se distancia ou se aproxima, em alguma medida, de outras variedades.

Visando desenvolver o trabalho investigativo aqui proposto, esta dissertação encontra-se assim estruturada: neste capítulo, abordarei o objeto da pesquisa, tomando por base não só a realidade linguística de Timor-Leste, em que há o contato linguístico do português com outras línguas, como também estudos já realizados sobre concordância e PSN em diversas variedades do português; no capítulo 2, discorrerei sobre aspectos teóricos a respeito da Faculdade da Linguagem e do PSN, abordando como ocorre, por exemplo, a aquisição de uma língua natural na perspectiva da GU; no capítulo 3, apresentarei os procedimentos metodológicos que

guiaram esta pesquisa e a constituição do *corpus*; finalmente, no capítulo 4, trarei os resultados quantitativos e farei uma análise dos dados do PTL e, por fim, serão enunciadas as considerações finais.

1.2 CONTATO LINGUÍSTICO DO PORTUGUÊS COM OUTRAS LÍNGUAS

A língua portuguesa partiu de uma realidade geográfica restrita à península ibérica, no século XV, para novos territórios por meio da expansão ultramarina de Portugal, o que permitiu que se fixasse em territórios na América, África e Ásia. Esse roteiro de disseminação linguística provocou, por conseguinte, o contato do português com línguas de grupos e famílias linguísticas diferentes. A saída definitiva de Portugal, no século XX, de todos os territórios ocupados a título de colônia deixou uma realidade linguística nesses territórios altamente modificada, em maior ou menor intensidade e em diferentes contextos, durante cerca de quinhentos anos:

> A partir de meados do século XV, a língua portuguesa, na esteira da expansão marítima de Portugal, sai de suas fronteiras europeias e se torna uma língua internacional, com falantes seus se estabelecendo em enclaves ao longo da costa ocidental e oriental do continente africano, alcançando Índia em 1948, a América em 1500, a China em 1513, o Timor em 1515 e o Japão em 1543. (FARACO, 2016, p. 57)

Esse alcance alçado pela língua portuguesa, no entanto, não costumou ser nem profundo nem altamente disseminado na maioria desses territórios. Além disso, não foi exatamente o que se entendia como língua portuguesa que chegou a vigorar nesses locais. Segundo Faraco (2016, p. 57), “o que funcionou como língua franca no cotidiano das navegações e do comércio foram antes o *pidgin* e, depois, as línguas crioulas de base portuguesa.” (2016, p. 57). Isso se justifica, em grande parte, porque o império mercantil português que se expandia não tinha propósito de estabelecer colônias de povoamento ou ocupar estes “novos” territórios. Antes disso, fez com que houvesse enclaves comerciais. Seus objetivos eram dois: estabelecer entrepostos para captar produtos localmente produzidos e ser portos de apoio para a rota marítima

que ia da Ásia até a Europa. Assim, a presença portuguesa nunca foi muito expressiva, muito menos nas regiões interioranas para além da Costa. Sendo a ocupação portuguesa realizada nesses termos, cabe que eu retome esta descrição panorâmica feita por Pereira (2007, p. 55):

> São muitas as referências ao uso da "língua portuguesa", em documentos oficiais e relatos de viagens dessa época. A designação de "língua portuguesa" que encontramos nesses textos encobre, por certo, variedades linguísticas muito diferentes: o português falado como língua materna pelos oficiais, soldados, comerciantes e marinheiros; o português escrito, tanto pelos portugueses como pelos outros europeus, em textos religiosos, cartas, relatos e documentos oficiais; o português formalmente aprendido, em casos pontuais, por alguns membros dos povos com que os portugueses entravam em contacto; as múltiplas variedades, mais ou menos bem-sucedidas, usadas nas tentativas de comunicação informal entre os povos africanos e asiáticos e os europeus; uma forma de linguagem veicular (*pidgin)* mais estável e regular, resultante do uso continuado dessas variedades; os próprios crioulos, que durante muito tempo não foram reconhecidos como línguas autónomas, mas como formas "corrompidas de português" e, finalmente, variedades aportuguesadas de crioulo a que um dos pioneiros dos estudos dos crioulos, Hugo Schuchardt, no século XIX, chamava "meio-crioulo" e que resultavam de tentativa, por parte de alguns falantes, da demarcação em relação aos " vulgares falantes de crioulo", através de uma aproximação em relação ao português.

Com o desenvolvimento do contato comercial entre portugueses e povos dos territórios alcançados, portanto, houve "a urgência do mútuo entendimento em contextos multilíngues [que] criou as condições para o surgimento... de *pidgin* de base portuguesa" (FARACO, 2016, p. 68) que, por sua vez, veio a se tornar, em alguns territórios, língua crioula. Para uma distinção entre *pidgin* e crioulo, Faraco (2016, p. 68-72) enuncia:

- *Pidgin*: linguagem emergencial que se caracteriza por léxico e gramática reduzidos, funcional apenas em situações restritas de comércio e trabalho e para a qual o contexto é fundamental para a inteligibilidade. Solução pragmática para um nível restrito de comunicação, marcado por eventos esporádicos, entre falantes de línguas diferentes e mutuamente incompreensíveis.

- *Crioulo*: língua nova resultante do desdobramento de *pidgin* que teve seus contextos de uso ampliados o suficiente para ser língua materna de comunidades estáveis (PARKVALL, 2012), ou resultante de

> complexo processo de hibridização linguística (MUFWENE, 2001). No segundo caso, são aproximadas línguas de substrato (da população local) e de superestrato (da potência europeia). Para o surgimento do crioulo, portanto, é necessário um contato extremo entre as línguas.

Ainda sobre o surgimento de línguas crioulas de base portuguesa, Pereira (2007, p. 15) observa que:

> Apesar das divergências quanto à definição do que é um crioulo, pelo menos num ponto os linguistas estão essencialmente de acordo: os crioulos distinguem-se das outras línguas pela rapidez de sua formação, em condições históricas fora do comum.

A autora também apresenta o necessário para identificar uma língua crioula de base portuguesa (Ibid., p. 47):

> Quando as unidades lexicais são, na sua maioria, reconhecidamente de origem portuguesa, embora, na sua estrutura, se rejam por regras fonológicas e morfológicas próprias, possam ter significados diferentes e impliquem condições sintáticas também diferentes.

De maneira geral, em crioulos de base portuguesa, o português influenciou destacadamente o componente lexical; já as línguas locais influenciaram outros componentes da gramática, como, por exemplo, o nível morfossintático. Cardoso, Hagemeijer e Alexandre (2015) agruparam esses crioulos em africanos e asiáticos e os subdividiram da seguinte forma, respectivamente: a) crioulos da Alta Guiné e crioulos do Golfo da Guiné e b) crioulos da Ásia meridional, Sudeste asiático e Ásia oriental. A partir da catalogação desses autores, apresento, no quadro 1, abaixo: a) os nomes das línguas crioulas, b) o subgrupo a que pertencem e c) a área geográfica onde são faladas:

Quadro 1 - Línguas crioulas de base portuguesa na África e na Ásia

Subgrupo	Nome	Área geográfica
Alta Guiné	Caboverdiano	Cabo Verde
	Kriyol	Guiné-Bissau
	Crioulo de Casamansa	Senegal
Golfo da Guiné	Santomé	Ilha de São Tomé
	Angolar	Ilha de São Tomé
	Principense	Ilha do Príncipe

	Fa d'ambô	Ilha de Annobón (Guiné Equatorial)
Ásia meridional	Crioulo de Diu	Índia
	Crioulo de Damão	Índia
	Crioulo de Bombaim	Índia
	Crioulo de Korlai	Índia
	Crioulo de Mangalor	Índia
	Crioulo de Cananor	Índia
	Crioulo de Mahé	Índia
	Crioulo de Cochim	Índia
	Crioulo de Negapatão	Índia
	Crioulo de Ceilão	Sri Lanka
Sudeste asiático	Crioulo de Malaca	Malásia
	Crioulo de Singapura	Singapura
	Crioulo de Batávia/Tugu	Indonésia
	Crioulo de Bidau	Timor-Leste
Ásia oriental	Crioulo de Macau	China
	Crioulo de Hong Kong	China

Fonte: Adaptado de Cardoso, Hagemeijer e Alexandre (2015)

A maioria das línguas crioulas apresentadas no quadro 1 não mais existem. Seus vestígios documentais foram usados pelos autores para a devida catalogação. Especialmente atenção merecem os crioulos caboverdiano, kriyol, santomé e principense, ainda presentes na África, bem como os de Malaca e Macau, que ainda existem na Ásia.

O crioulo de Malaca teve forte expressão na região de todo o sudeste asiático. Albuquerque (2011) indica que os crioulos de Malaca, Macau e Bidau influenciaram consideravelmente a realidade linguística de Timor-Leste, bem como a língua Tétum, que será abordada adiante. O crioulo de Bidau, conforme afirmam Cardoso, Hagemeijer e Alexandre (2015, p. 673), "foi registado pela última vez na década de 50 do século XX (Thomaz 1985; Baxter 1990) e, ao que tudo indica, desde então terá deixado de ser falado." A título de exemplo das características de um crioulo português, transcrevo a seguir um trecho documental do crioulo de Malaca, extraído de Cardoso, Hagemeijer e Alexandre (2015, p. 683), composto de duas estrofes de uma cantiga de Malaca (Rêgo 1998 [1942]):

Pastorinho berde,	*Passarinho verde,*
Más bêrde di rico flor;	*Mais verde do que a rica flor;*
Bai lebá êste chito,	*Leva este recado,*
Dá com eu sa amor.	*E dá-o ao meu amor.*
Pastorinho berde,	*Passarinhos verdes,*
Um ramo santá dôs dôs;	*Sentam-se num ramo aos pares;*
Eu nádi morrê lonzi,	*Eu não morrerei longe,*
Eu logo more perto bôs.	*Eu morrerei perto de ti.*

Analisando essas estrofes, Cardoso, Hagemeijer e Alexandre (2015, p. 683) verificam:

> Aqui encontramos alguns traços importantes do crioulo de Malaca, como sejam linha 4 a marcação do objeto pela prep. *com* (na ortografia atual, *ku*) ou a utilização de um marcador gen. pós-nominal *sa* (do ptg. *sua*): ex. *eu sa amor* (linha 4). Este crioulo faz uso produtivo da reduplicação com funções icónicas, nas quais se incluem a indicação de pluralidade e intensidade, mas também, como se vê em *dôs dôs* (‘aos pares’, i.e. ‘dois a dois’), a de distributividade. A primeira estrofe contém ainda uma construção comparativa na qual se verifica que o padrão é marcado por uma prep. locativa com valor abl., linha 2 a prep. *di* em *di rico flor*. Nos dois últimos versos, observa-se uma dicotomia nos marcadores pré-verbais de fut., com uma forma positiva *logo* e outra negativa *nádi* (do ptg. *não há-de*).

É, portanto, evidente a influência do português no léxico do trecho reproduzido, ao passo que a estrutura e escrita fogem das características dessa língua, visto serem advindas das línguas locais, ou autóctones. Sendo assim, passo agora a apresentar um panorama da situação linguística de Timor-Leste, tendo em mente o contato linguístico do português com outras línguas.

1.2.1 A realidade linguística de Timor-Leste

A ilha de Timor localiza-se no mar do sudeste asiático e é um território multilíngue (HULL, 2001b). Ali, durante os séculos XV e XVI, foi ponto de intensas rotas de navegação e comércio que ligavam a Europa à Ásia. Como expoente das campanhas de navegação, Portugal se estabeleceu em diversos pontos da região asiática, como Goa (Índia), Macau (China), a ilha de Malaca (Malásia) e a ilha de Timor. A administração portuguesa lançava mão de línguas francas para se comunicar com os povos dessa grande área, que eram habitadas por falantes de diversas línguas,

vindas de diferentes troncos linguísticos. Em Timor, as línguas catalogadas (cerca de dezesseis) parecem vir de, pelo menos, duas famílias: a austronésia e a papuásica (ALBUQUERQUE, 2011). Com o tempo, desenvolveram-se crioulos de base lexical portuguesa na Ásia, conforme já apresentado no quadro 1. De acordo com Albuquerque (2011), o Crioulo Português de Macau, o Crioulo Português de Malaca e o Crioulo Português de Bidau eram falados no chamado Timor Português dos séculos XVII a XIX. Pela pouca quantidade de portugueses residentes em Timor, os crioulos, mais do que a língua portuguesa, influenciaram e deram vários empréstimos às línguas da ilha. O Tétum, língua franca local e atualmente oficial de Timor-Leste, sofreu grandes influências dos crioulos de base lexical portuguesa, especialmente do Crioulo Português de Bidau, falado na capital, Díli, a partir do século XVIII.

Mais recentemente, a parte leste da ilha de Timor sofreu com novas invasões estrangeiras. Primeiro, o Japão ocupou a ilha durante a Segunda Guerra Mundial. Após isso, em 1975, a Indonésia ocupou a parte que até então era conhecida como Timor Português, na tentativa de anexá-la ao seu território. Acredito ser digno de nota o período do domínio indonésio (que durou até a independência de Timor-Leste, com a intervenção das Nações Unidas em 1999) porque, além de ser marcado pela violência e brutalidade, esse domínio também se utilizou de comandos para regular o uso das línguas: desincentivou o tétum e proibiu o ensino e o uso do português, estabelecendo a língua indonésia como obrigatória. Segundo Albuquerque (2011, p. 35),

> Na área linguística, a dominação indonésia adotou uma política linguística de *indonesização* da população timorense, assim como o planejamento linguístico indonésio mostrou-se ter extrema eficiência. Nos vinte e quatro anos em território timorense, a Indonésia conseguiu: diminuir o pouco que restava do uso da língua portuguesa pela população timorense banindo totalmente seu uso, sendo utilizada somente pelos revolucionários como *língua de resistência*; a língua Tetun como possui um caráter cultural e identitário para a nação timorense foi reduzido seu uso lentamente no dia-a-dia da população; o *bahasa indonesia* (língua oficial e nacional da Indonésia) foi implantado de maneira intensa por meio de um grande fluxo de professores e materiais didáticos, assim como da importação de diversos itens da cultura material indonésia como livros, revistas, músicas, programas de rádio e televisão, produtos de consumo em geral, entre outros.

Apenas em 2002, quando foi reconhecida como nação, a República Democrática de Timor-Leste tomou iniciativa para reavivar o tétum (na variedade mais conhecida pela população) como língua nacional e oficial e o português como língua

oficial do país, promovendo o seu ensino e amplo uso. Para isso, o governo timorense se valeu de acordos de cooperação com países como Brasil e Portugal, que enviam professores, materiais didáticos e outros profissionais das mais diversas áreas, bem como formam timorenses para atuar como professores de língua portuguesa em Timor, a exemplo do que acontece na UNILAB, localizada nos estados do Ceará e da Bahia (CARIOCA, 2016). Albuquerque (2011, p. 38) afirma que

> a única iniciativa de planejamento linguístico é em relação ao ensino de língua portuguesa no nível pré-secundário da rede de ensino público de Timor-Leste. Há uma preocupação de natureza política em se implantar a língua portuguesa nos primeiros anos de ensino educacional.

Atualmente, percebe-se a grande relevância do tétum e a tentativa persistente de inserção do português em Timor-Leste. Mesmo assim, o português ainda é falado com fluência por menos de 10% da população, sendo uma Língua Estrangeira (LE)[1], nas palavras de Paraíso (2019), visto que é ainda aprendido, na maioria das vezes, após o período crítico da aquisição da linguagem:

> O Tetun Prasa é falado como L1 por 18% da população, enquanto 82% da população leste-timorense possui alguma fluência nele, pois é a língua que funciona como língua franca, ou seja, língua de comunicação entre os diferentes grupos etnolinguísticos que possuem L1 distintas. (...) Digno de nota, é que a parcela da população leste-timorense falante de língua portuguesa, assim como a falante de língua inglesa, vem crescendo significativamente nos últimos anos, por causa da política linguística adotada pelo Estado. (ALBUQUERQUE, 2011, p. 47-48)

Albuquerque (2011, p. 53) ainda afirma que, “apesar de uma quantidade considerável de empréstimos lusófonos, a língua Tetun é de origem austronésia, mais especificamente da família Malaio-Polinésia.” Sobre a classificação dessa língua, figuram estudiosos definindo-a como crioulo ou *pidgin* (HULL, 2001b). Outros consideram-na uma variedade das várias línguas Tétum (Tétum Díli, Tétum Terik e outras) que existem na ilha e que, devido a um intenso contato com outras línguas, passou por processos de reestruturação gramatical (ALBUQUERQUE, 2011). Como há certa variação na própria denominação da língua, utilizarei a forma “tétum”, que é

[1] Quanto à nomenclatura aqui apresentada, Língua Estrangeira (LE), permanece como usada por Paraíso (2019). No segundo capítulo desta dissertação, a nomenclatura que utilizarei será a de Língua Não Materna (LNM) para me referir ao PTL, contrapondo-se tão somente ao conceito de Língua Materna (LM).

a grafia oficial em Timor-Leste. Além disso, é importante especificar a que variedade do tétum estou me referindo, pois, segundo Sofia Deus (2011, p. 235), o tétum é:

> língua da qual também existem variedades muito distintas. O *Tétum-térique* é uma variedade da língua hoje falada essencialmente na região de Soibada, sendo também percepcionada como uma variedade utilizada especialmente em contextos literários e mais eruditos. Outra das variedades desta língua é o *Tétum ocidental* ou *"Belunês"*, o qual é de influência malaia e é falado na região oriental do Timor indonésio (portanto, perto da fronteira com Timor-Leste) e ainda nos distritos de Balibó e Suai. Finalmente, existe ainda a variedade conhecida como *Tétum-Díli* (designação equivalente às de *Tétum-praça* ou de *Tétum-oficial*), que é língua franca nacional, sendo oficial a par do Português e língua materna dos falantes do distrito de Díli. Neste estudo, é esta última variedade que nos propomos investigar, pela sua nítida relevância no panorama linguístico, social, político e educacional de Timor-Leste.

Assim como a autora propõe, a língua tétum a que me refiro em toda essa pesquisa é o Tétum-Díli, língua nacional e oficial de Timor-Leste[2]. No que se refere às influências do português na língua, elas dizem respeito ao léxico, à morfologia e à sintaxe. No domínio lexical, seguem exemplos, como:

bazar: 'mercado'
kolega: 'palavra usada para se referir a amigos, ou pessoas que se tenha uma intimidade e que sejam da mesma idade'
argoliña: 'brinco'
mestri: 'professor de escola'
dosenti: 'professor universitário'
alunu/a: 'estudante do ensino fundamental ou médio'
estudanti: 'estudante universitário'
valor: 'resultados de exames escolares'.
karreta: 'carro'

(ALBUQUERQUE, 2011, p. 75)

Itens lexicais mais recentemente trazidos do português dão conta de termos específicos e têm sentido tal qual a sonoridade que nos remete ao português:

[2] Dentre as línguas locais de Timor-Leste, o Tétum é a mais conhecida e estudada. Daí o fato de recorrer às características dessa língua nos exemplos desta dissertação, assim como de fazer uma apresentação de seu léxico e de sua estrutura morfossintática. Não levantarei, portanto, a hipótese de que só e somente o Tétum é predeterminante na configuração morfossintática do PTL, haja vista o intenso contexto multilinguístico em que se insere essa variedade não-europeia do português.

> ekolojía, biolojía, linguístika, istória, siénsia, sientífiku, administrasaun, eleisaun, votu, governu, prezidenti, deputadu, embaixador.
>
> (ALBUQUERQUE, 2011, p. 75)

Na questão da morfologia de flexão verbal, não há variação desinencial do verbo no tétum (DEUS, 2009):

Quadro 2 - Paradigma de flexão verbal do PE e do Tétum

Português Europeu	Tétum-Díli
Eu falo	Ha'u *ko'alia*
Tu fal*as*	Ita / O *ko'alia*
Ele/Ela/Você fal*a*	Nia *ko'alia*
Nós fal*amos*	Ita / Ami *ko'alia*
Vós fal*ais*	Imi *ko'alia*
Eles/Elas/Vocês fala*m*	Sira *ko'alia*

Fonte: Deus (2011, p. 238)

As informações de modo e tempo virão acompanhadas em forma de marcadores (lexemas). Abaixo seguem exemplos nos quais esses lexemas (destacados em negrito) expressam: a) perfectivo (**tiha**) em (1), b) tempo-aspecto perfeito (**tiha-ona**) em (2), c) aspecto iterativo (**nafatin**) em (3) e futuro (**sei**) em (3) e (4). Esses são apenas alguns marcadores lexicais do Tétum, conforme apresentados por Albuquerque (2011, p. 104-109) num esboço gramatical dessa língua:

(1) haree **tiha** ema lubun boot, zezus sa ba foho ida
ver PERF pessoa quantidade grande Jesus subir para montanha IND
'Quando viu a multidão, Jesus subiu à montanha.'

(2) mestri haruka **tiha-ona** labarik sira tuur ba rai.
professor mandar PF criança PL sentar para terra
'O professor mandou as crianças sentarem no chão.'

(3) efeitu sira husi krize **sei** sente **nafatin** iha timoor rai.
efeito PL ABL crise FUT sentir ITE LOC timor terra
'Os efeitos da crise serão sentidos continuamente em terras leste-timorenses'

(4) hau **sei** halo kafee mai ami.
1sg FUT fazer café para 1exc
'Eu farei café para nós'

No que diz respeito à sintaxe, o Tétum tem ordem canônica SVO e a constituição dos sintagmas segue a parataxe:

> (O Tétum) é uma língua SVO desprovida de morfologia de flexão verbal, quer de número, quer de pessoa ou de tempo, ou seja, as formas verbais são completamente invariáveis. Como tal, a ordem de palavras numa frase tétum é de extrema importância, visto ser a posição das mesmas que determinará a sua função. (DEUS, 2011, p. 235)

A ordem dos constituintes permanece Sujeito-Verbo para todos os tipos de verbo. Outro aspecto morfossintático é que, ao contrário do que ocorre no PE, língua de ensino dos timorenses, os sujeitos devem ser realizados foneticamente no tétum, ou seja, a posição de sujeito deve ser ocupada por sujeitos plenos (cf. (5a), sendo agramaticais orações com sujeito nulo (cf. (5b)). Realmente, não há como identificar esse sujeito pela morfologia de flexão verbal, uma vez que ela é pobre (cf. quadro 2):

(5)a. **Ha'u** ko'alia lian Indonésio.
Eu falo Indonésio.
b. ***pro** Ko'alia lian Indonésio.

(DEUS, 2009, p. 46-47)

Não obstante, é possível que sujeitos nulos sejam legitimados no tétum, sob algumas restrições. São elas: a) contexto de oração encaixada "co-indexados com um operador-wh" (cf. (6)), sendo proibido um quantificador como antecedente na oração principal (cf. (7)) e b) sujeitos nulos expletivos (não-argumentais) (cf. (8)):

(6) a. Se maka hanoin katak maka matenek, foti liman.
Quem é que pensar COMP é que esperto levantar mão
"Quem pensa que é[3] *pro* esperto, levante a mão.

[3] Em tétum, os verbos que expressam cópula não são realizados foneticamente.

b. Se maka hanoin katak han bolu, foti liman.
Quem é que pensar COMP comer bolo levantar mão
“Quem pensa que *pro* come bolo, levante a mão.” (quem = suj. encaixado)

(7) a. La iha ema ida maka hetan livru sira-ne’ebé maka *(nia) halakon.
não haver pessoa uma que encontrar livro PL qu- é que perder
“*Ninguém* encontrou os livros que perdeu.”

b. Ema hotu-hotu hetan livru sira-ne’ebé maka *(sira) halakon.
pessoa todo encontrar livro PL qu- é que perder
“*Todos* encontraram os livros que perderam.”

(8) a. Foufoun, iha de’it kolega timoroan uitoan, maibé agora iha barak ona.
Princípio haver só colega timorense um pouco mas agora haver muito já
“No princípio, *havia* poucos colegas timorenses, mas agora *há* muitos.”
b. Tarde.
Tarde
“*pro* era tarde.”

(DEUS, 2011, p. 236-237)

Importa dizer aqui que é sobre este último aspecto morfossintático, ou seja, relativo ao preenchimento da posição sujeito, que centrarei minha atenção nesta dissertação, pois ainda são escassos estudos sobre a morfossintaxe do PTL (BRITO, 2010). As pesquisas disponíveis foram feitas recentemente, isto é, a partir da independência definitiva de Timor-Leste do domínio estrangeiro, que ocorreu apenas em 2002. Em geral, os estudos linguísticos abordam a língua tétum e outras línguas faladas na ilha. (ALBUQUERQUE, 2011, 2012):

> Há poucos trabalhos sobre esta variedade do português, e os que existem nenhum se dedicou ao estudo da morfossintaxe, existindo apenas breves comentários, sendo o primeiro elaborado por Vasconcelos (1970 [1901]:184). Posteriormente, há Thomaz (1985), reproduzido em Thomaz (2002:153); Carvalho (2002/2003); Brito e Bastos (2007) e Albuquerque (2011a). (ALBUQUERQUE, 2012, p. 4)

Como uma pauta importante é o ensino da língua portuguesa na ilha, há estudos sobre essa temática (SILVA; TELES, 2015; GUIMARÃES, 2017), evidenciando que há muitas influências das línguas locais no português que lá se aprende/fala (SANTOS, 2009; DEUS, 2009; MARTINS, 2016).

Há pesquisas na área da aquisição/aprendizagem do português por indivíduos de Timor-Leste que afirmam que “o sistema verbal do português europeu é de difícil aquisição para os falantes timorenses”. (MARTINS, 2016, p. 70). Assim sendo, espero contribuir com os estudos linguísticos sobre a morfossintaxe do PTL, a fim de verificar, por exemplo, como se comportam os sujeitos, levando em conta se há ou não possíveis interferências do multilinguismo nessa variedade do português. Um ponto a destacar é que os dados analisados nesta pesquisa não são provenientes da língua falada, mas da língua escrita, em específico, de redações produzidas por alunos timorenses que visavam ao seu ingresso na UNILAB. Esses dados escritos por alunos escolarizados poderão proporcionar uma visão geral sobre as características do PTL, especialmente em contexto de maior monitoramento da língua. Como não estão no período crítico de aquisição da linguagem, a escrita dos alunos deve mostrar como a língua já adquirida, sobretudo, o tétum, assim como as demais línguas locais, interfere na língua aprendida, o português.

Visando traçar um panorama dos estudos já realizados em outras variedades do português, abordarei, na próxima seção, o preenchimento da posição de sujeito, a fim de ampliar o entendimento para uma questão central deste estudo: o PTL teria mais sujeitos nulos, sendo mais próximo da variedade europeia do português, ou mais sujeitos plenos?

1.3 CARACTERIZAÇÃO DO OBJETO DE ESTUDO

1.3.1 Incursões no Parâmetro do Sujeito Nulo[4]

O preenchimento da posição de sujeito nas línguas naturais é previsto na GU porque essa posição é comum a todas as línguas. Chomsky (1981) apresentou o Princípio de Projeção Estendido, que estabelece que essa posição é sempre projetada em todos os níveis de representação em todas as línguas particulares. O que difere essas línguas tem a ver com preenchimento dessa posição por uma categoria plena (sujeito pleno) ou por uma categoria nula (sujeito nulo, nomeadamente *pro*), culminando assim na proposta do PSN, que está associado a um conjunto de propriedades, como a omissão do sujeito, a inversão livre do sujeito em frases simples, riqueza da morfologia de flexão verbal, dentre outras (CHOMSKY, 1982).

A rica natureza de AGR é um dos fatores que permitem a omissão do sujeito na sentença, uma vez que a concordância possibilita retomar, pelos morfemas número-pessoais, as informações relacionadas ao sujeito (RIZZI, 1986). Uma das línguas com AGR rico que atende às propriedades do PSN é o espanhol, que junto ao italiano e o PE, são consideradas línguas de sujeito nulo prototípicas (RIZZI, 1985, 1986; SILVA, 2004). O português brasileiro (PB), por outro lado, se distancia um pouco em relação ao PE. Nele, temos um enfraquecimento da morfologia de flexão verbal e uma maior tendência ao preenchimento da posição sujeito por sujeitos plenos (DUARTE, 1995). Ao passo que o PE e o PB se distanciam, variedades africanas do português, de maneira geral, tendem a se aproximar mais da variedade europeia. Entretanto, estudos recentes mostram que é possível encontrar variedades em uma posição intermediária entre o PB e o PE, como é o caso do português de Moçambique (PM) (VERÍSSIMO, 2018).

A relação estabelecida inicialmente entre morfologia de flexão verbal rica e legitimação e identificação de sujeitos nulos (RIZZI, 1982) foi revisitada e ampliada para outros três tipos de sistemas de sujeito nulo, a saber: a) língua de sujeito nulo radical (ex.: chinês), b) língua de sujeito nulo expletivo (ex.: alemão); c) línguas de

[4] Nesta subseção, faço apenas uma breve introdução sobre o PSN e, no capítulo seguinte, mais especificamente, na seção 2.2, discorro sobre esse parâmetro no âmbito da Teoria de Princípios e Parâmetros (CHOMSKY, 1981 e seguintes).

sujeito nulo parcial (ex.: PB) (HOLMBERG, NAYUDU, SHEEHAN, 2009). Sendo assim, em vez de só definir uma língua como sendo de sujeito nulo, é importante definir em que sistema de sujeito nulo ela se insere. O estatuto do PSN será abordado com mais detalhes no segundo capítulo desta dissertação.

Ao contrário de línguas de sujeito nulo, há aquelas que têm uma morfologia de flexão verbal fraca e não legitimam sujeitos nulos, como o inglês e o tétum.

Sobre a concordância em variedades do português, estudos mostram que o PE apresenta uma regra de concordância verbal semicategórica, independentemente do grau de escolaridade dos portugueses (BRANDÃO; VIEIRA, 2012), o que implica considerar que seu paradigma de flexão verbal é rico, ao contrário do que se observa no PB, em que a regra é variável, sinalizando o enfraquecimento de sua morfologia de flexão verbal. Um aspecto a pontuar é que, na língua falada dessa última variedade, há mais ausência de concordância verbal quando comparado ao que ocorre na língua escrita, sobretudo, quando se observam os alunos que têm maior grau de escolarização (MAGALHÃES, 2000). No que se refere a algumas variedades africanas do português, a concordância é enfraquecida quando se trata de pessoas com baixa escolaridade, ao contrário do que se observa com falantes com alta escolaridade (MOTA; MIGUEL; MENDES, 2012).

1.3.2 parâmetro do sujeito nulo em variedades do português

Nesta subseção, abordarei o PSN em diferentes variedades do português, a fim de que possa traçar um panorama de possíveis semelhanças e/ou diferenças entre elas.

Classificado na literatura como uma língua de sujeito nulo consistente (RAPOSO, 1992; DUARTE, 1995; SILVA, 2004) ao lado de outros sistemas de línguas de sujeito nulo (HOLMBERG, NAYUDU; SHEEHAN, 2009), o PE enquadra-se nessa classificação devido à riqueza de sua morfologia de flexão verbal, que é capaz de identificar os sujeitos nulos, conforme já previsto inicialmente por Taraldsen (1978), Chomsky (1981) e Rizzi (1982, 1986). Em (9), os morfemas número-pessoais no PE identificam os sujeitos nulos (nomeadamente, ***pro*** ou ***ø***) com traços de primeira pessoa do singular (c. (9a)) e de primeira pessoa do plural (cf. (9b)). Tal situação não

é verificada no inglês, cuja morfologia de flexão verbal é pobre, ocasionando a obrigatoriedade de sujeitos plenos (cf. (10)):

(9)a. ***pro*** Jogo.
b. ***pro*** Joga*mos*.

(10)a. **I** play.
b. **We** play.

Silva (2004) ainda observa que, no PE, o sujeito nulo referencial é a opção mais frequente dos falantes (cf. (11)), enquanto no PB, mesmo em contextos referenciais, os falantes tendem a realizar o sujeito (cf. (12)):

(11) a. Comemos o bolo.
b. Eles pensam ter*mos* comido o bolo.
(RAPOSO, 1992, p. 482)

(12) a. Eu estou com fome.
a'. *Ø estou com fome.[5] (KATO, 1999, p. 1)

Contrariamente ao PE, estudos mostram que a morfologia de flexão verbal do PB está passando por um enfraquecimento, ocasionando uma reorganização em várias áreas de sua gramática, dentre elas a grande frequência de sujeitos plenos, ao invés de nulos (DUARTE, 1995, 2000; GALVES, 2001; SILVA, 2004; KATO; DUARTE, 2014). Duarte (1995), por exemplo, atestou como o princípio "evite pronome" vem se perdendo na língua falada do PB. Esse enfraquecimento se evidencia por apresentar duas formas zero: uma na segunda pessoa do singular e outra na segunda pessoa do plural, conforme quadro a seguir:

[5] Acredito não ser possível assumir categoricamente que (12a') é agramatical no PB, mas sim que é a opção menos realizada entre os falantes. Silva ressalva que, enquanto Kato (1999) considerou a ocorrência agramatical, outros autores, como Figueiredo Silva (1996), trazem como gramaticais construções de primeira pessoa em orações livres, como "Comprei um carro ontem"; de igual maneira, é comum, ao menos em alguns dialetos do PB, a forma "*Tô* com fome", com o sujeito nulo.

Quadro 3 - Paradigma de flexão verbal do PB e do PE

Português Brasileiro	Português Europeu
*(eu) cant**o***	*(eu) cant**o***
----	*(tu) canta**s***
*você cant**a*** *ele cant**a***	*você cant**a*** *ele cant**a***
*(nós) canta**mos***	*(nós) canta**mos***
----	----
*vocês canta**m*** *eles canta**m***	*vocês canta**m*** *eles canta**m*** ‘

Fonte: Galves (2001, p. 103)[6]

Embora a morfologia esteja se enfraquecendo no PB, sujeitos nulos ainda são produzidos nessa língua, conforme exemplificado na oração em (13), extraída do *corpus* de uma pesquisa já realizada por mim (OLIVEIRA JUNIOR, 2018):

(13) “Ao decorrer do dia, *pro* podemos observar crianças e jovens indo a caminho da escola.”

A oração acima apresenta um sujeito nulo (*pro*) de primeira pessoa do plural ocupando a posição do sujeito, pois o morfema número-pessoal *-mos* identifica-o. Contrariamente, no inglês, o sujeito nulo seria impossível nessa oração como em (14a[7]), devido à obrigatoriedade da realização fonética do sujeito na posição sujeito (cf. (5b)):

(14) a. *During the day, *pro* can observe children and young ones on the way to school.”

b. “During the day, *we* can observe children and young ones on the way to school.”

Apesar de o *corpus* da pesquisa que realizei (OLIVEIRA JUNIOR, 2018) ter sido elaborado a partir de redações produzidas por brasileiros que já haviam concluído o nível médio e que prestavam exames para seu ingresso no ensino superior,

[6] Galves (2001, p. 124) também observa que há dialetos do PB que “mostram contraste apenas entre a primeira pessoa do singular e todas as outras: *eu canto/você, nós, eles canta*”.

[7] Tradução minha do exemplo apresentado em (13).

verifiquei, naquela ocasião, que os resultados vão na direção dos de Duarte (1995, 2000), que analisou a língua falada: embora sujeitos nulos ocorram, há alta produtividade de sujeitos plenos mesmo que se trate de um contexto de escrita mais monitorada. Com base nisso, surge um questionamento: estariam os dados escritos do PTL apresentando um comportamento semelhante aos do PB no sentido de que a frequência de sujeitos plenos seja alta?

Em seu trabalho pioneiro sobre o PSN no PB, Duarte (1995) verificou uma mudança no paradigma flexional dessa língua, que está se enfraquecendo, conforme já mencionado: afastando-se da variedade europeia do português, o falante brasileiro tende a realizar foneticamente os sujeitos pronominais e referenciais nas sentenças, de modo que o PB estaria perdendo o princípio “Evite pronome”. Para tanto, a autora trabalhou com entrevistas de falantes cariocas advindas do projeto NURC, bem como entrevistas e gravações de programas de TV. Nessa pesquisa, Duarte encontrou sujeitos plenos (cf. (15) e (16)) e nulos (cf. (17) e (18)) sendo produzidos.

(15) Essas queixas *elas* existem, é claro.
(16) *Ela* só deve se defragar quando...
(17) O Japão é outro mundo. *pro* É [uma cultura completamente diferente].
(18) *pro* Eram razoavelmente pagas. (ref. “as professoras”)

(DUARTE, 1995, p. 37-38)

Em seus resultados sobre a distribuição de sujeitos de referência definida na fala espontânea, a autora contabilizou uma grande diferença, alcançando 71% de produtividade de sujeitos plenos, em relação a 29% de sujeitos nulos. Isso se verificou também no que diz respeito aos verbos na 1ª pessoa, tanto no singular, como no plural. A única marca categórica de sujeito nulo em seus dados ocorreu na 2ª pessoa do plural. Para ela,

> O português brasileiro perdeu a propriedade que caracteriza as línguas de sujeito nulo do grupo *pro-drop* por força do enfraquecimento da flexão, responsável pela identificação da categoria vazia sujeito em línguas que apresentam uma morfologia verbal suficientemente “rica” para tal processo. (DUARTE, 1995, p. 141)

No entanto, a autora não descarta a produção de sujeitos nulos no PB, pois a mudança do PB de língua *pro-drop* para língua *não-pro-drop* em decorrência da perda

da especificação gramatical de alguns morfemas número-pessoais ainda não havia sido implementada:

> Essa perda, no entanto, não se reflete no uso da língua como mudança concluída. O PB atual convive com um sistema agonizante, em que ainda se refletem as características de *pro-drop,* e um sistema em desenvolvimento, em que a "riqueza funcional" perdida já não permite a identificação de *pro.* (DUARTE, 1995, p. 141-142)

Em outras palavras, Duarte (1995) seguiu na direção de propor um processo de mudança de um sistema para o outro, de *pro-drop* para *não-pro-drop,* e afirmou que o sujeito pleno estaria ganhando na produção dos referenciais em PB. Como veremos mais à frente, essa proposta foi revisitada pela autora, que atualmente considera o PB uma língua de sujeito nulo parcial (KATO; DUARTE, 2014a, 2014b), tomando por base os quatro sistemas de línguas de sujeito nulo propostos por Homlberg, Nayudu e Sheehan (2009).

Por sua vez, comparando as características do sujeito nulo no PB, no finlandês e no chinês, Modesto (2004) argumenta que as três línguas são *pro-drop,* nas quais a recuperação do sujeito nulo é possível sem a marcação conferida pela flexão verbal. Segundo ele, os sujeitos dessas línguas são tópicos elididos por estarem inseridos em uma cadeia discursiva. Centrando a atenção nos sujeitos nulos referenciais de sentenças encaixadas, foi defendida a ideia de que as línguas em questão eram, igualmente, orientadas para o discurso, conforme pode ser observado nos exemplos nessas línguas:

(19) a. [O amigo do Feco$_2$]$_1$ disse que $e_{1/*2/*3}$ ganhou a competição.

b. [Veljeni$_2$ vaimo]$_1$ oli ninn eloinen, ettei $e_{1/*2/*3}$ voinut nukkua.
irmão-GEN esposa estava tão feliz que-não podia dormir
'A esposa do meu irmão estava tão feliz que não conseguia dormir.'

c. [Zhangsan$_2$ fanwen de ren]$_1$ xiwang $e_{1/*2/*3}$ neng ying.
Zhangsan visitar DE pessoa esperar pode vencer.
'A pessoa que Zhangsan visitou esperava que pudesse ganhar.'

(MODESTO, 2004, 128)

Nos exemplos acima, temos a retomada do tópico inteiro (entre colchetes) como sujeito da segunda da oração encaixada em português (cf. (19a)), finlandês (cf. (19b)) e chinês (cf. (19c)), não sendo possível uma interpretação diversa desse sujeito encaixado nem para um elemento interno ao tópico ou qualquer referente externo à frase.

No caso do chinês, em particular, se observado apenas o paradigma de flexão verbal nessa língua, pode-se dizer que seja uma língua *não-pro-drop*, mas, quando levados em conta aspectos discursivos, trata-se de uma língua *pro-drop*, pois o sujeito nulo da encaixada está ligado a um antecedente já mencionado no discurso com o qual mantém correferência. Para explicar como são legitimados sujeitos nulos em línguas de orientação para o discurso, Modesto (2004) propõe o seguinte:

> A ideia principal é a de que em línguas orientadas para o discurso o núcleo F(ocus)P sempre é gerado (em toda oração) como um traço OCC (=EPP) e, então, algum constituinte é necessariamente movido para a posição de SPEC desse núcleo e interpretado como tópico contrastivo. (Ibid. p. 135)

Explicando essa ideia, o autor nos apresenta a seguinte oração:

(20) [FP o Feco$_1$ [IP t_1 convenceu a Dani$_2$ [CP que [FP ele$_1$ [IP t_1 pode se eleger]]]]]

(MODESTO, 2004, p. 136)

Nessa oração, “o sujeito encaixado e o sujeito matriz formam uma cadeia de tópicos e o primeiro pode ser deletado em PF” (MODESTO, 2004, p. 136). Assim, há relação entre o movimento e a interpretação do constituinte como antecedente do sujeito nulo em PB, uma análise que se estende ao chinês e ao finlandês.

Em uma pesquisa desenvolvida por Holmberg, Nayudu e Sheehan (2009), os autores verificam que, embora permitam o sujeito nulo, o PB, o finlandês e o marata apresentam restrições para a produção desses sujeitos, o que os leva, a partir da proposta de quatro sistemas de línguas de sujeito nulo, classificarem essas línguas como línguas de sujeito nulo parcial. Sobre o PB, por exemplo, sujeitos não argumentais e de referência genérica são omitidos. A respeito dos segundos, os autores apresentam a seguinte distinção entre o PB (cf. (21) e (22)) e PE (cf. (23) e (24)):

(21) É assim que faz o doce. (PB)
(22) Nesse hotel não pode entrar na piscina bêbado. (PB)

(23) É assim que se faz o doce. (PE)
(24) Nesse hotel não se pode entrar na piscina bêbado. (PE)
(HOLMBERG, NAYUDU & SHEEHAN, 2009, p. 62-63)

O preenchimento com o clítico “se” em PE é necessário para a leitura genérica do sujeito, o que não ocorre em PB. Além disso, nos casos de nulos em encaixadas, como nos exemplos abaixo, os autores defendem que o sujeito nulo é controlado por um antecedente em uma sentença superior por meio da correferência:

(25) O João$_1$ disse que (ele$_1$) tinha comprado uma casa.
(26) Os meninos$_1$ ficavam contentes quando (eles$_1$) tinham um dia de folga.
(HOLMBERG, NAYUDU & SHEEHAN, 2009, p. 65)

As conclusões a que os autores chegaram é de que, nas três línguas em análise (PB, finlandês e maraca), sujeitos nulos ocorrem em contextos não permitidos em línguas *não-pro-drop,* como o inglês e o francês, podendo apresentar uma cadeia de sujeitos nulos encabeçada e controlada por um tópico não-nulo no discurso, bem como sujeitos nulos de interpretação genérica ou impessoal.

Seguindo a proposta de Holmberg, Nayudu e Sheehan (2009), Kato e Duarte (2014) fazem importantes considerações sobre o PSN no PB e, revisitando a proposta de Duarte (1995), propõem que não é o princípio ‘Evite Pronome’ em PB que está em jogo, mas o princípio ‘Evite pronomes não-referenciais’, característico de línguas de sujeito nulo parcial. Nesse sentido, o PB constitui-se como uma língua de sujeito nulo que pode não vir a se tornar uma língua *não-pro-drop,* bem como não é uma língua de sujeito nulo radical como o chinês, uma língua orientada para o discurso. Segundo as autoras, o PB apresenta “propriedades de uma gramática estável caracterizável em termos de restrições semânticas de referencialidade e de propriedades morfossintáticas” (Ibid., p. 17). Por exemplo, assim como Galves (2001) já salientava, sujeitos mais referenciais tendem a ser realizados foneticamente, tomando por base a hierarquia de referencialidade, em que a referencialidade é considerada “uma

propriedade gradiente do mais específico para o menos específico, sendo os menos específicos aqueles itens não dotados do traço [+humano]. (Ibid., p. 17). De igual modo, sujeitos menos referenciais (os expletivos, por exemplo) são nulos:

(32) $\emptyset_{expl}$ chove muito nessas florestas.
(33) Essas florestas chovem muito.
(34) $\emptyset_{expl}$ Estourou o pneu do Hamilton.
(35) [O Hamilton]i estourou [o pneu t_i].

(KATO & DUARTE, 2014, p. 12)

Por fim, elas afirmam que, no caso dos sujeitos referenciais, “o sistema permite uma variação entre o sujeito nulo e o sujeito expresso” (p. 18). Além disso, por serem os pronomes sujeitos do PB fracos, é também possível seu apagamento (Ibid. p. 18).

Vale trazermos para esta discussão o trabalho desenvolvido Souza et. al. (2010) sobre a distribuição de sujeitos plenos (cf. (27)) e nulos (cf. (28)) referenciais em textos escritos produzidos por estudantes do Ensino Fundamental (5ª à 8ª série) da cidade de Florianópolis:

(27) *Eu* fiquei furiosa porque faltava um dia para começar minhas aulas.
(28) Todos os finais de semana *Ø* vou para a casa de minha vó.

(SOUZA et al., 2010, p. 98-99)

Os resultados gerais da pesquisa mostram 37% de ocorrências de sujeitos nulos e 63% de sujeitos plenos, um resultado que dialoga com o que foi obtido por Duarte (1995) para a língua falada do PB no século XX. Nesse estudo, uma variável extralinguística com impacto considerável foi a escolaridade, visto que mais sujeitos nulos foram produzidos à medida que a escolaridade aumentava:

> Nota-se que os alunos das séries mais baixas preenchem mais o sujeito pronominal, o que pode ser um indicativo de que, na escola, ainda seja aprendido o princípio “evite o pronome” (cf. DUARTE, 1995). Portanto, quanto mais tempo na escola, menos o aluno preenche o sujeito pronominal, corroborando também resultados estatísticos de Costa (2003) e de Brustolin (2009). (SOUZA et. al., 2010, p. 102)

Algo que chama atenção nessa pesquisa diz respeito ao uso de sujeitos nulos de primeira pessoa em uma oração imediatamente posterior a outra em que o sujeito é a forma pronominal *a gente*:

(29) Depois, *agente*[8] *saímos* do shopping e *Ø fomos* no parque de diversões.
(30) *A gente fomos* pra praia. *Ø Voltamos* só a noite e o resto dos dias *Ø ficamos* brincando.
(31) *Agente* se *arrumamos* e *Ø fomos*.

(SOUZA et. al., 2010, p. 104)

Em linhas gerais, as autoras concluem que "a realização [de sujeitos plenos] se revelou significativa possivelmente porque o PB está perdendo a possibilidade de recuperar o referente através do elemento de concordância" (Ibid., p. 105), bem como para evitar a ambiguidade entre pronomes como "a gente" e "você" quando identificados com traço de terceira pessoa do singular, visto que os verbos acompanhados desses pronomes são conjugados com as marcas número-pessoais da terceira pessoa. À semelhança de Duarte (1995), defendem que "o Português do Brasil está em fase de transição de uma língua pro-drop para uma língua não pro-drop" (Ibid., p. 105).

Diante do exposto sobre o PB e o PE, centraremos nossa atenção a partir de agora no PSN em variedades africanas do português. De antemão, vale referirmos que, a partir de um levantamento bibliográfico nessas variedades sobre a morfologia de flexão verbal e o PSN, verifiquei, até o momento, que ainda são escassas as pesquisas.

Mota, Miguel e Mendes (2012) afirmam que, em variedades africanas do português, como o português de Angola e o português de São Tomé, falantes com maior grau de escolarização tendem a um uso linguístico mais próximo da variedade europeia do português. Especialmente no que se refere à concordância, as autoras elencaram os princípios mais relevantes para o afastamento ou aproximação da concordância em relação à norma do PE, quais sejam:

[8] A grafia dos exemplos citados nesta dissertação sempre mantém a forma como foi publicada no texto da referência.

> No verbo, (i) a ausência de marcas de concordância é quantitativamente irrelevante em PE e tipologicamente limitada – o PE caracteriza-se globalmente pela reiteração das marcas morfológicas de concordância, pela «concordância dissociada» (cf. Costa & Silva 2006), ou seja, por uma regra de concordância quase categórica; (ii) nas restantes variedades nacionais, há evidências no sentido de uma tendência para o aumento da presença de marcadores de concordância, embora ocorram casos de ausência em percentagens ainda relevantes, (iii) havendo alguma diversidade de padrões entre variedades não europeias; (iv) os factores extralinguísticos com maior relevância (nas zonas urbanas estudadas, em geral) são o nível de escolaridade e a idade (geralmente indissociáveis). (MOTA, MIGUEL, MENDES, 2012, p. 166)

Assim, no caso de sujeitos escolarizados, espera-se que haja uma maior riqueza do paradigma de flexão verbal e, consequentemente, maior licenciamento de sujeitos nulos.

Baseada em dados produzidos por estudantes universitários moçambicanos, Gonçalves (2007) fez um levantamento do que chamou de “erros” de aquisição do português como L2 em Moçambique, o que a levou a argumentar a favor do surgimento de uma “nova gramática” no português de Moçambique (PM), que era desenvolvida a partir de uma dificuldade de adaptar completamente o *input* da L2 para a devida reorganização gramatical. Dentre os tipos de “erros” verificados pela pesquisadora, destacam-se os chamados erros de concordância verbal, sobretudo, no que se refere à concordância de número: “a população universitária auscultada apenas comete erros de concordância em número, não se tendo registado nenhum erro de concordância em pessoa” (GONÇALVES, 2007, p. 71). Observem-se os seguintes exemplos extraídos dessa pesquisa:

(36) Antigamente [os jovens] **era** aconselhados pelos mais velhos. (PE: eram)
(37) [As doenças, a fome e a morte] **é** outra tragédia. (PE: são)
(38) [−] **existiam** [uma outra família] que se opunha a esse relacionamento. (PE: existia)

(GONÇALVES, 2007, p. 73)

Mesmo apresentando esses casos de “erros”, a autora conclui que não são muito expressivos: “neste caso não há nenhum tipo de erro que sobressaia particularmente, oscilando a sua percentagem cerca de 15% e 33%.” (Ibid., p. 73). E ainda,

> Os erros de concordância verbal parecem ter ainda carácter disperso, podendo admitir-se que a sua ocorrência é devida à falta de exposição dos falantes a evidências robustas (negativas ou positivas explícitas) sobre a língua-alvo. (Ibid., p. 74)

Deste modo, é possível concluir da pesquisa de Gonçalves que há uma aquisição diferente da esperada do português como L2 em Moçambique, atribuída, em parte à pouca robustez das evidências linguísticas às quais o falante tem contato. Sendo assim, essa variedade diverge um pouco em relação ao PE quando observados os usos linguísticos de falantes universitários.

Outra pesquisa sobre a morfologia de flexão verbal no PM é desenvolvida por Pissurno (2017) que se volta à concordância verbal de terceira pessoa do plural na língua falada do PM. A autora esclarece que, embora o português seja língua oficial, a realidade linguística de Moçambique é múltipla: o português convive com mais de 20 línguas da família bantu, destacando-se o makhuwa, com 25% de falantes, que é expressivamente mais utilizado que o português, com 11% de falantes. Em regiões como a capital Maputo, o PM tem maior expressão, atingindo quase 43% dos falantes. A pesquisadora ainda destaca que o bilinguismo está presente em praticamente todas as regiões do país, sendo que, em áreas mais rurais, o PM tem menor uso na comunicação cotidiana.

Para o desenvolvimento de sua pesquisa, foram entrevistados 35 informantes em Maputo, que foram divididos em três níveis de instrução: Fundamental, Médio e Superior. Das 2.353 ocorrências verificadas, a concordância verbal padrão teve um alto percentual, 96,8% dos casos, de acordo com o quadro abaixo:

Quadro 4 - Distribuição dos dados com e sem marca verbal de 3ª pessoa plural no PM

	OCORRÊNCIAS	PERCENTUAL
Concordância padrão	2278/2353	96.8%
Concordância não padrão	75/2353	3.2%

Fonte: Pissurno (2017, p. 155)

Além disso, os casos em que não houve a concordância estão mais relacionados aos informantes que falam mais as línguas locais do que o português.

Por sua vez, a variante escolaridade analisada revelou que aqueles com Ensino Superior também utilizaram menos concordância não padrão em comparação aos demais grupos, indo na direção do PE:

> os resultados aqui apresentados demonstram que os dados da concordância verbal na variedade moçambicana, em termos quantitativos, à semelhança do que ocorre com as variedades europeias, apresentam larga preferência pela marcação padrão. (PISSURNO, 2017, p. 209)

Há, no entanto, situações em que o PM apresenta marcação distinta do PE e também do PB na questão da concordância verbal de 3ª pessoa. Em vista disso, a autora propõe um *continuum*[9] entre essas línguas, sendo que, no PB, a regra de concordância é variável, seguida pelo PM e, por último, o PE em que a essa regra apresenta-se como semicategórica, seguindo a proposta de Labov (2003) (cf. BRANDÃO; VIEIRA, 2012).

Centrando sua atenção no PSN, em seu trabalho sobre os sujeitos anafóricos de 3ª pessoa no PM, Santos (2010) nos relembra que essa variedade é falada por 39% da população e se constitui como L2, convivendo com a L1, que sempre é uma língua da família bantu. Por meio de entrevistas realizadas com 4 informantes de nível superior, constatou a alta produtividade de sujeitos nulos, com percentual de 72%. Segundo ela, a variedade moçambicana se aproxima consideravelmente da europeia, ao passo que se distancia da brasileira, cujo percentual de nulos atingiu apenas 37%, de acordo com a pesquisa de Barbosa, Duarte e Kato (2001). Ademais, quando analisadas as variáveis acessibilidade e animacidade do antecedente, a variedade moçambicana esteve mais próxima da europeia e mais distante da brasileira:

> (...) não só permitiu a confirmação da hipótese de que o PB apresenta um comportamento diferente do PE e do PM como pôde se confirmar a suposição de que a variedade moçambicana do português apresenta semelhanças com a língua alvo, o PE, em que não se verificam restrições ao apagamento do sujeito estando ou não o antecedente sintaticamente acessível. (SANTOS, 2010, p. 38)

À semelhança do que verificou Santos (2010), Oliveira e Santos (2007), ao analisarem o PSN no português de Angola (PA) e ao compararem seus resultados com o PE e com o PB, concluíram que essa variedade africana do português é uma língua *pro-drop*, tal como dito para o PE: “o PA está próximo à variante PE: o não-preenchimento da posição de sujeito é o mais recorrente nos dados” (Ibid., p. 97). No quadro 5, é apresentado o total de ocorrências de pronomes plenos e nulos extraídos

[9] Cf. Brandão e Vieira (2018) também defendem esse *continuum* entre as variedades.

das 369 sentenças analisadas, tomando por base as pessoas do paradigma pronominal:

Quadro 5 - Número de ocorrências das formas pronominais na posição de sujeito no PA

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

Fonte: Oliveira e Santos (2007, p. 97)

De acordo com o quadro 5, é possível perceber que houve percentual expressivo do sujeito nulo, representando cerca de 75% de todas as ocorrências de sujeito mapeadas. Assumindo a hipótese da "hierarquia de referencialidade" proposta por Cyrino, Duarte e Kato (2000) para o PB, os autores observam que também no PA sujeitos mais referenciais representados, por exemplo, pela 1ª pessoa do singular e do plural, tendem a ser plenos. Conforme já dito, apesar da existência desses sujeitos, os autores concluem que o PA tem preferência por sujeitos nulos, o que o caracteriza como uma língua *pro-drop,* indo na direção do PE.

Em se tratando do português de Cabo Verde (PCV), Lopes e Campos (2015), ao compará-lo com o PB, analisaram a distribuição de sujeitos nulos e plenos nessas variedades.

Sendo adquirido como L2 em Cabo Verde, o português vive em situação de bilinguismo, e, como hipótese, eles consideraram que seria "bem possível que a expressão pronominal não fosse completamente similar à de Portugal." (Ibid. p. 321) em virtude da influência do crioulo caboverdiano (CCV), crioulo de base lexical portuguesa que existe há séculos e que é língua materna, mas não oficial, da população, sendo, portanto, a L1 da maioria dos falantes; nessa língua, sujeitos plenos são obrigatórios e a morfologia de flexão verbal, fraca (BAPTISTA, 2002; PRATAS, 2002, 2004).

Segundo Lopes e Campos (2015), “no PCV, (há) uma tendência para o preenchimento dos sujeitos das orações principais, sobretudo, da primeira pessoa do singular — eu, em contraste com a primeira do plural — nós, menos preenchida” (Ibid. p. 333). Os exemplos, a seguir, ilustram isso:

(39) 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.

(40) 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?

(LOPES; CAMPOS, 2015, p. 334)

Para além de outras considerações realizadas, os autores concluem que existem semelhanças entre o PCV e o PB, como “o clítico *se* com o seu uso como pronome *default*, tal qual ocorre no PB vernacular, e o licenciamento de pronomes tônicos como objetos diretos.” (Ibid. p. 341). No caso do PCV, o uso de sujeitos plenos justifica-se, em grande parte, por conta do contexto de bilinguismo em que essa variedade se encontra.

Destaque-se também aqui a pesquisa de Malta (2019) sobre o PSN com base na comparação de três variedades africanas do português por meio da análise de textos escritos por falantes do PA, do português de São Tomé e Príncipe (PST) e do português de Guiné-Bissau (PGB). Ao longo desse estudo, foi analisada a natureza da morfologia verbal e sua implicação para a produção de sujeitos nulos e, para tanto, foram extraídos dados de redações escritas por falantes com nível médio que se submeteram a um vestibular para ingresso na UNILAB localizada na cidade de Redenção-CE.

Durante a análise, a pesquisadora verificou no PA o seguinte quantitativo de sujeitos nulos e plenos:

Quadro 6 - Quantitativo geral de sujeitos plenos e nulos nos *corpora* do PA

	Sujeitos nulos	Sujeitos plenos
Português de Angola	1.925,2/3.753 51,3%	1.827,7/3.753 48,7%

Fonte: Malta (2019, p. 134)

O resultado mostra que o PA tem sujeitos nulos ligeiramente mais frequentes que os plenos, apenas 2,6 pontos percentuais, sendo assim uma variação estável. Ela acentua que a morfologia de concordância verbal é rica no PA e que isso faz com que a regra de concordância nessa variedade seja semicategórica em direção ao PE.

Em sua análise dos textos escritos por falantes do português de São Tomé e Príncipe (PST), Malta (2019) revela que: "o português é falado por mais de 95% na ilha de São Tomé, sendo usada como L1 por mais de 50% da população em detrimento do santomé, língua crioula dessa ilha." (Ibid., p. 63). O quadro, a seguir, apresenta o quantitativo de sujeitos plenos e nulos encontrados nos textos analisados pela autora:

Quadro 7 - Quantitativo geral de sujeitos plenos e nulos nos *corpora* do PST

	Sujeitos nulos	Sujeitos plenos
Português de São Tomé e Príncipe	1.391,78/2.626 53%	1.234,22/2.626 47%

Fonte: Malta (2019, p. 159)

O quantitativo obtido revela um índice de realização do sujeito nulo similar ao que a autora encontrou no PA, superando os sujeitos plenos por poucos pontos percentuais. Esse resultado aponta para uma variação estável no uso das duas formas e pode ser reflexo da regra variável de concordância verbal já verificada por Brandão e Vieira (2012) na língua falada dos santomenses.

Por sua vez, ao analisar o PGB, Malta (2019) afirma que, em Guiné-Bissau, o multilinguismo também é presente, mas com uma distinção importante: a principal língua utilizada na comunicação é o crioulo guineense, o kriol. Considerando toda a população do país, Malta (Ibid., p. 58) afirma que "o PGB é atualmente falado por cerca de 1,5 milhões de guineenses, ou seja, cerca de 57% da população." Há uma considerável carga histórica e cultural no crioulo, que é mantido como língua da

comunicação por boa parte dos falantes, sendo assim o português, por ser língua oficial, fica mais restrita ao domínio da escola e da administração. Visando quantificar o percentual de uso de sujeitos nulos e plenos nessa variedade, a autora obtém os seguintes resultados:

Quadro 8 - Quantitativo geral de sujeitos plenos e nulos nos *corpora* do PGB

	Sujeitos nulos	Sujeitos plenos
Português de Guiné-Bissau	827,2 20%	3.308,8 80%

Fonte: Malta (2019, p. 181)

No PGB, observa-se que a produção de nulos é menos expressiva do que no PA e no PST. A autora considera que, diferindo em realidade linguística tanto do Brasil como de Portugal, o multilinguismo interfere sobremaneira nos dados observados: "para o PGB, parece haver evidências fortes da influência do crioulo guineense sobre os dados" (Ibid., p. 198) e, por isso, essa variedade não se aproxima do PE, mesmo nos contextos de escolarização, muito provavelmente pela carência de políticas de ensino que direcionem melhor a abordagem do português nesse país. Além disso, a concordância verbal é variável no PGB e, sobre o uso de sujeitos nulos, verifica que

> são licenciados nos dados escritos (...) pela presença de um morfema flexional no verbo, pela referência semântica da P4 e P6 (primeira e terceira pessoas do plural), como também pelo traço semântico [-animado] do sujeito. Dessa forma, podemos perceber (que) o efeito expressivo do favorecimento do sujeito nulo está condicionado às orações encaixadas, enquanto nos contextos de outras orações a tendência é de desfavorecimento do sujeito nulo (MALTA, 2019, p. 199).

Em linhas gerais, após a retomada de estudos em diferentes variedades do português, observo que algumas das variedades africanas do português tendem a um uso mais próximo do PE no que se refere à tendência a um alto percentual de uso da concordância verbal e ao uso de sujeitos nulos, valendo ressaltar que já há propostas que integram algumas variedades africanas, como é o caso do PM, em uma posição intermediária entre o PB e o PE (VERÍSSIMO, 2017). Diante disso, ergue-se uma das questões deste trabalho: estaria o PTL mais próximo de que variedades no que se refere à concordância e ao PSN?

1.4 OBJETIVOS

1.4.1 Geral

Mapear a distribuição de sujeitos nulos e plenos em um *corpus* de língua escrita do português de Timor-Leste (PTL), observando se a morfologia de flexão verbal interfere na produção desses sujeitos.

1.4.2 Específicos

1 - Analisar a natureza da morfologia de flexão verbal no *corpus* e suas possíveis implicações para o uso dos sujeitos nulos;
2 - Observar se há uma possível influência do contexto de multilinguismo em Timor-Leste para o preenchimento da posição sujeito nos dados de escrita do PTL;
3 - Discutir os resultados obtidos com os de outros estudos já realizados sobre o PSN em variedades do português, levando em conta com qual(quais) variedades o PTL evidenciaria maior proximidade ou afastamento.

1.5 HIPÓTESES

1.5.1 Geral

O contexto multilíngue do Timor-Leste pode ter influenciado a configuração morfossintática do português falado ali. Sendo assim, os dados escritos de timorenses escolarizados devem revelar uma alta produtividade de sujeitos plenos nos mais variados contextos, bem como um paradigma de flexão verbal enfraquecido, diferentemente do PE.

1.5.2 Específicas

1 - O PTL pode evidenciar enfraquecimento em sua morfologia de flexão verbal em decorrência da fraqueza da morfologia de flexão verbal das línguas locais, como o tétum; consequentemente, pode haver baixa produtividade de sujeitos nulos pelo fato de a morfologia de flexão verbal não ser capaz de identificar os traços número-pessoais do sujeito;

2 - O contexto multilíngue pode estar exercendo influência sobre aspectos morfossintáticas do PTL, uma Língua Não-Materna, tais como: maior produtividade de sujeitos plenos e uso de verbos sem marcas flexionais;

3 - O PTL pode apresentar maior proximidade com variedades não-europeias do português em contextos multilíngues em que o português convive com línguas *não-pro-drop* adquiridas como línguas maternas e bastante utilizadas no cotidiano dos falantes.

1.6 CONCLUSÃO DO CAPÍTULO

Neste primeiro capítulo, após a contextualização realizada visando à delimitação do objeto de estudo, bem como a apresentação de algumas pesquisas sobre concordância e PSN em variedades do português, observo que há tendência a um distanciamento do PE das demais variedades quando o português não é a língua majoritária e está inserido em contexto multilíngue. Em Timor-Leste, essas características estão presentes e, portanto, os resultados a serem apresentados e discutidos mais adiante podem evidenciar também esse distanciamento. Mesmo no caso de PB, em que o português é a língua majoritária do território, as características morfossintáticas já permitem classificá-lo como língua de sistema de sujeito nulo diferente do PE. E o que dizer do PTL?

2 BASE TEÓRICA DA PESQUISA

2.1 A TEORIA GERATIVA: PRESSUPOSTOS BÁSICOS

2.1.1 A hipótese inatista

A Gramática Gerativa é um programa de investigação científica que entende a linguagem como capacidade cognitiva e biológica natural do homem. Este termo, cunhado por seu pesquisador pioneiro, Noam Chomsky[10], tem sido usado para se referir a várias pesquisas que buscam especificamente descrever e explicar (1) a natureza, (2) a origem e (3) o uso da linguagem humana. Diante disso, seguem as questões centrais da teoria gerativa levantadas por Chomsky (1971):

1. O que, de fato, é o conhecimento linguístico que temos em nossa mente?
2. Como uma criança adquire esse conhecimento linguístico, em vista do fato de que em poucos anos ela consegue se comunicar utilizando todas as complexas estruturas linguísticas que temos à disposição?
3. Como o conhecimento linguístico presente em nossa mente é utilizado para produzir as mais diversas expressões das línguas naturais?
4. Como o cérebro, o centro da atividade mental e cognitiva, produz e lida com esse conhecimento linguístico?

Entendida como parte da cognição humana, a capacidade da linguagem (denominada de Faculdade da Linguagem (FL)) é um órgão biológico, natural. Isso significa afirmar que todos os humanos já nascem com essa estrutura mental responsável pela linguagem. Entender essa hipótese do inatismo linguístico é, sem dúvida, uma resposta a um questionamento sobre como as crianças adquirem o conhecimento linguístico (*core grammar*) em um tempo relativamente curto (por volta

[10] Noam Chomsky nasceu em 1928, em Filadélfia, Pensilvânia, EUA. Formou-se na Universidade da Pensilvânia, tendo obtido ali o título de Doutor e foi pesquisador assistente na Universidade de Harvard. Lecionou por mais de 40 anos no MIT (*Massachusetts Institute of Technology*), EUA. Conhecido como pai da linguística moderna, ele deu grande impulso aos estudos formais da linguagem, numa perspectiva racionalista, contrária ao que propõe o behaviorismo.

dos cinco anos) e a partir de um *input* tão limitado e deteriorado. Em outras palavras, no genoma humano, existem “materiais genéticos inatos que se destinam especificamente à aquisição e ao uso da linguagem.” (KENEDY, 2013, p. 74).

Chomsky (1998, p. 18) ressalta ainda que a capacidade humana da linguagem é “em grande parte, responsável pelo fato de, sozinhos no universo biológico, os seres humanos terem uma diversidade e evolução cultural de alguma complexidade e riqueza” se comparados a qualquer outra espécie de vida, estando a uma absurda distância delas nessa perspectiva. No entanto, teorizar sobre a linguagem passa pela necessidade de conceituar “língua”.

Na perspectiva gerativista, dois conceitos de língua são discutidos, a saber: língua-E e língua-I. A língua-I, por um lado, é o conhecimento gramatical internalizado do falante de uma ou mais línguas. Ela reúne “um conjunto de capacidades e habilidades mentais que fazem um indivíduo produzir e compreender um número potencialmente infinito de expressões linguísticas da língua de seu ambiente” (KENEDY, 2013, p. 34).

Por outro lado, entende-se como Língua-E tudo aquilo que é externo à estrutura mental da linguagem. Trata-se de uma construção histórica e social que é elaborada por diferentes grupos de indivíduos que têm fatores de identificação, como parentais, geográficos, culturais, e cuja composição revela contingências e características específicas desses grupos. Ela se revela em um código, que costumamos denominar de língua, a exemplo da *língua portuguesa, língua italiana, língua tétum.* Esses códigos são altamente complexos e sua composição está para além de fatores linguísticos internos. Recebe, por exemplo, interferência do contexto, com a criação de novos vocábulos, alterações e simplificações históricas. Recebe também interferência política, desde a sua denominação, padronização, popularização e ensino. Sofre, ainda, interferência do meio social e do próprio indivíduo, que atuam sobre o seu código, ampliando, modificando e ressignificando a este. O seu uso é guiado por fatores ideológicos e, portanto, é constantemente modificado.

Em suma, a língua-E está para o desempenho, performance, assim como a língua-I, para a competência. Em outras palavras, a segunda diz respeito ao conhecimento linguístico internalizado do falante que independe de instrução formal, ao passo que a primeira diz respeito ao uso desse conhecimento em diferentes contextos cotidianos (KATO, 1995).

Para adquirir uma língua natural, existe um estado inicial (S_0) que se denomina Gramática Universal (GU), já presente na mente-cérebro do falante e, portanto, inata. Sobre essa gramática, Chomsky (1980, p. 28) afirma que é um

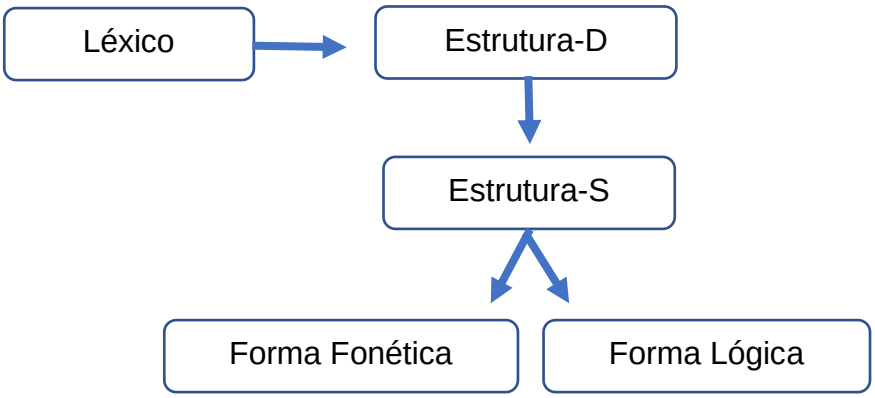
> Sistema de princípios, condições e regras que são elementos ou propriedades de todas as línguas humanas, não por mero acaso, mas por necessidade – quero dizer, é claro, necessidade biológica, e não lógica. Então, pode-se dizer que GU expressa "a essência da linguagem humana". GU será invariante entre os humanos. GU especificará o que a aprendizagem da língua deve alcançar, se for bem sucedida. (CHOMSKY, 1980, p. 28)

Desde a década de 80, com o Modelo de Princípios e Parâmetros (P&P) (CHOMSKY, 1981), defende-se que já estão previstas na GU propriedades universais, invariantes, de todas as línguas (a saber: os Princípios). A título de exemplo, há o **Princípio de Projeção Estendido** formulado por Chomsky (1982), que estabelece que, em todas as línguas naturais, a posição de sujeito é sempre projetada em orações finitas em todos os níveis de representação que, segundo Chomsky (1965), são: a Estrutura-D (estrutura profunda), a Estrutura-S (estrutura superficial) e os níveis de interface, a forma fonética (PF, do inglês *Phonetical Form*) e a forma lógica (LF, do inglês *Logical Form*). Sob essa perspectiva teórica, Silva (2004, p. 89) resume as funções desses níveis:

> a) a Estrutura-D (DS) em que são representadas de forma transparente as relações locais de subcategorização e as funções lógico-gramaticais que se estabelecem entre os itens lexicais já formados no léxico e que são inseridos na sintaxe; b) a Estrutura-S (SS) que corresponde à representação concreta da estrutura frásica no que concerne à sua hierarquização; c) a Forma Fonética (PF) que faz a interface com os sistemas articulatório e perceptual e d) a Forma Lógica (LF) que faz a interface com os sistemas conceitual e intencional.

Percebe-se que o léxico é o material linguístico primário que será manipulado no decorrer da Estrutura-D e da Estrutura-S, pelas quais as expressões linguísticas serão harmoniosamente geradas e enviadas aos dois sistemas de interface, responsáveis pela adequação e representação sonora e conceitual dos enunciados. Veja-se o esquema dos níveis de representação:

Figura 1 - Níveis de representação, de acordo com Princípios e parâmetros



Fonte: Chomsky (1981, p. 17)

Ao lado dos Princípios, propriedades específicas, variantes (a saber: os Parâmetros) também compõem a GU, cujos valores binários ((+) ou (-)) serão fixados pela criança durante o processo de aquisição da linguagem após sua exposição aos dados linguísticos primários (*input*). Um exemplo de parâmetro tem a ver com as categorias que podem ocupar a posição de sujeito nas línguas naturais: há línguas em que essa posição deve ser preenchida obrigatoriamente por sujeitos realizados foneticamente, não permitindo, portanto, sujeitos nulos como é o caso do inglês (cf. (41)) (línguas *não-pro-drop*) e línguas que permitem sujeitos não-realizados foneticamente a exemplo do espanhol (línguas *pro-drop*) (cf. (42)):

(41) a. I play soccer.
Eu jogo futebol
b. *Play soccer.
Jogo futebol

(42) a. Yo voy a su casa.
Eu vou à sua casa
b. Voy a su casa.
Vou à sua casa.

A partir de P&P, mais especificamente, os linguistas gerativistas vêm se ocupando a não só descrever o conhecimento internalizado do falante, mas também

a explicá-lo, tomando por base a estrutura cognitiva responsável pela linguagem e os mecanismos ativados na produção e compreensão das estruturas frásicas. Nesse sentido, Chomsky (1998 [1968], p. 24) defende que “uma genuína teoria da linguagem humana tem que satisfazer duas condições: ‘adequação descritiva’ e ‘adequação explicativa’”.

Com base nas considerações feitas até aqui, é possível observar, do ponto de vista epistemológico, uma mudança causada nas pesquisas sobre linguagem que vinham sendo desenvolvidas no início do século XX. A proposta da linguagem como capacidade inata do ser humano e da existência de um órgão mental responsável pela aquisição e constituição da competência linguística revela que o foco do estudo da linguagem está na mente do indivíduo, e não nos dados linguísticos a que ele está exposto.

Os estudos do Estruturalismo americano acabaram por investigar, sob a ótica do behaviorismo, a língua como um sistema de regras produzidos por uma determinada comunidade de fala cujos dados linguísticos são observáveis e, portanto, finitos. Opondo-se a esse modelo de análise e às teorias do comportamento, Chomsky (1975, p. 26) defende que “só podemos aprender algo de significativo sobre essa natureza [da linguagem] investigando a capacidade cognitiva do organismo, e essa investigação não permitirá ‘boas experiências’”. Nesse sentido, uma disciplina que desconsidere esse componente cognitivo

> estará condenada em princípio à investigação de assuntos periféricos, tais como a velocidade e o alcance da aquisição de informações, a relação entre a ordenação dos reforços e a força da resposta, o controle do comportamento, e similares. A disciplina em questão pode continuar indefinidamente a reunir informações sobre tais assuntos, mas cabe questionar o interesse ou o objetivo de tais esforços. (...) Algumas realizações intelectuais, tais como a aprendizagem da língua, estão estritamente dentro da capacidade biologicamente determinada. Para tais empreendimentos estamos “especialmente planejados”, de sorte que estruturas cognitivas de grande complexidade e interesse se desenvolvem com certa rapidez e pouco – se é que existe – esforço consciente. (CHOMSKY, 1975, p. 26)

E ainda, argumentando a favor de estudos que contemplem a cognição, diz Chomsky (1975, p. 27):

> No caso da cognição humana, é ao estudo das estruturas cognitivas básicas, dentro de sua capacidade cognitiva, seu desenvolvimento e aplicação, que se deve dar prioridade, creio eu, se quisermos obter um conhecimento verdadeiro sobre a mente e seu funcionamento.

Para além da existência do aparato biológico, a distinção entre *competência* e *desempenho* colabora bastante para identificar e resolver os chamados “problemas de linguagem”. Por meio dessa perspectiva, os profissionais envolvidos no diagnóstico de distúrbios da comunicação podem mais efetivamente definir tais dificuldades como sendo de origem cognitiva/mental específica da linguagem ou de outra natureza. Neste segundo caso, temos um problema de *desempenho*. Ele pode estar relacionado com questões motoras, sensoriais, fonológicas, físicas, emocionais, dentre tantas outras, mas não com a linguagem em si.

Uma contribuição valiosíssima tem a ver com o espaço que se deu às línguas de sinais como línguas naturais das comunidades surdas. O indivíduo surdo, portanto, não tem nenhum “problema de linguagem” e consegue, tal qual qualquer outro ser humano, adquirir e utilizar-se da língua no seu dia a dia, pois nasce com a GU e, no processo de aquisição, também fixará todos os valores dos parâmetros necessários à configuração de sua gramática internalizada. Também usará esse conhecimento linguístico por meio da modalidade gestual como expressão de sua interface articulatória/perceptual (ou sensório/motora, cf. Chomsky, 1993).

Observo, desse modo, que os ideais histórico-políticos sobre supremacia, prevalência de línguas e culturas caem todos por terra. Sujeitos de sociedades ágrafas ou inseridas no mundo da escrita digital; falantes de países pobres ou altamente ricos; línguas de filiação linguística conhecidas de longa data e línguas mais recentes, frutos do contato linguístico: todos eles e todas elas provêm exatamente do mesmo componente biológico, que está presente em todos os seres humanos por excelência. Verceze (2009, p. 98) afirma que

> Chomsky sintetiza essa idéia no livro *Conhecimento de Linguagem (Knowledge of Language)* (1986) procurando mostrar que os indivíduos, mesmo com contatos limitados com o mundo, são capazes de conhecer tanto e reproduzir pela linguagem este mundo.

2.1.2 A Faculdade da Linguagem

Chomsky (1965, 1975) defende que os seres humanos possuem uma Faculdade da Linguagem (FL), isto é, um órgão mental encarregado dos aspectos

cognitivos relacionados à linguagem humana. Essa proposta veio da necessidade de se teorizar a respeito da capacidade humana de gerar inúmeros enunciados linguísticos, muitos deles complexos, numa infinidade de combinações a partir de um número limitado de itens (FODOR, 1983; HAUSER; CHOMSKY; FITCH, 2002).

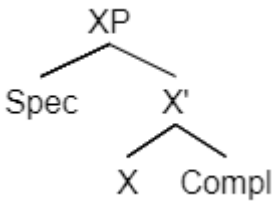
Seguindo a proposta de a cognição humana ser organizada em diferentes módulos que interagem entre si, a FL também é modular, sendo composta por diferentes módulos com princípios próprios e que interagem entre si no momento da geração dos enunciados linguísticos. Eles são responsáveis por seu funcionamento e nos permitem vislumbrar a complexidade da geração desses enunciados. Dentre esses módulos, há, por exemplo, a Teoria X-barra, a Teoria temática, a Teoria do Caso, entre outros.

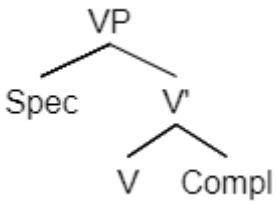
A Teoria X-barra procura dar conta de como os sintagmas, com base em sua natureza, são estruturados, de como as relações se estabelecem dentro deles e de como eles se hierarquizam para formar a sentença (JACKENDOFF, 1977; CHOMSKY, 1986; ABNEY, 1987). A Teoria estabelece que cada categoria sintagmática tem uma categoria lexical como seu núcleo. Para um núcleo verbal (V), nominal (N), adjetival (A) ou preposicional (P) teremos um VP, NP, AP e PP, respectivamente, ou seja, para um sintagma XP, X é o seu núcleo. Trata-se, portanto, do princípio da endocentricidade (WILLIAMS, 1981a). O núcleo X é responsável por determinar as relações internas do sintagma que ocorrem em dois níveis: o nível X', a projeção intermediária, e o nível XP, a projeção máxima do núcleo. Nessa perspectiva, uma hierarquia entre categorias é estabelecida (CHOMSKY, 1981).

A partir dos sintagmas de núcleo lexical, mais abaixo na hierarquia, tem-se também os sintagmas de categoria funcional, mais acima, como, por exemplo: IP (*Inflectional Phrase*) e CP (*Complementizer Phrase*). Essas categorias funcionais: "(i) possuem uma (e uma só) posição de especificador; (ii) formam classes fechadas (e restritas) de elementos; (iii) não possuem o valor semântico normalmente associado às categorias lexicais e (iv) subcategorizam obrigatoriamente um (e um só) complemento" (RAPOSO, 1992, p. 209).

Na representação arbórea em (45 a.), exemplifica-se a estrutura hierarquizada dos sintagmas proposta pela Teoria X-barra.

(45) a. b.





Em (45 b.) temos a representação de um sintagma verbal (VP). A Teoria X-Barra agrupa os constituintes sempre aos pares, de modo que o núcleo lexical V subcategoriza um sintagma, que é seu complemento (Compl), para juntos projetarem o nível intermediário V'. Este, por sua vez, une-se a um especificador (Spec) para atingir a projeção máxima do sintagma (VP). Essa hierarquia é comum a todas as línguas, mas a posição do núcleo em relação ao seu complemento pode variar, o que culminou no chamado Parâmetro de Ordenação do Núcleo (STOWELL, 1981). As línguas, de acordo com esse parâmetro, podem ser de núcleo inicial, como o português, com ordem Verbo-Objeto, ou de núcleo final, como o japonês, cuja ordem é Objeto-Verbo.

A Teoria Temática, por sua vez, é o módulo onde são estabelecidas as relações entre predicadores e argumentos. Também chamada de Teoria-θ, essa teoria define que um predicador X selecionará a quantidade e a natureza dos argumentos necessários à composição de sua estrutura temática na Estrutura-D (CHOMSKY, 1981). Observemos como se constitui a grelha temática do predicador verbal *comprar*:

(46) a. João comprou uma galinha

b. comprar: categoria [−N, +V]
nº de argumentos [−, −]
c-seleção [DP, DP]
s-seleção [Agente, Tema]

Conforme apresentado em (46), o predicador V *comprar* seleciona dois argumentos. Seguindo os critérios de c-seleção (seleção categorial), que estabelece a categoria dos argumentos, os dois são DP. Seguindo os critérios de s-seleção (seleção semântica), por sua vez, que estabelece o papel temático dos argumentos,

o argumento externo (que ocupa a posição de sujeito na sintaxe) terá papel-θ de Agente e o outro (que ocupa a posição de complemento na sintaxe), de Tema. Sendo assim, a Teoria-θ tem as funções de explicar quais elementos são capazes de atribuir e receber papel temático, qual o nível sintático e as posições em que se dá a atribuição e o recebimento de papéis temáticos e os princípios que regulam essa atribuição (CHOMSKY, 1981, 1986; MIOTO; SILVA; LOPES, 2018).

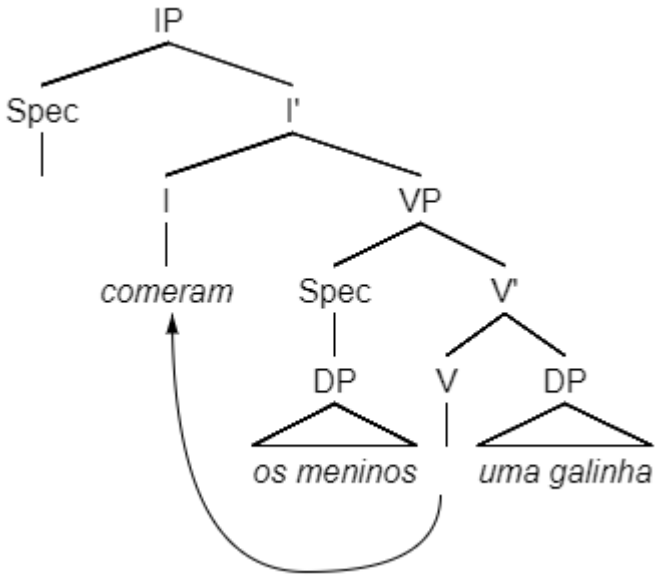
De acordo com Baker (1997), há outros papéis temáticos além de Agente e Tema, como Causativo, Experienciador, Benefactivo, Instrumento, Locativo e outros. Os núcleos que agirão como predicadores trazem consigo, do Léxico, as propriedades de sua grelha temática e a atribuição de papéis temáticos, por sua vez, ocorre na DS.

A Teoria do Caso, por sua vez, precisa dar conta de quais são os Casos abstratos, que categorias os atribuem, quais sintagmas os recebem e como esses Casos são atribuídos.

Os Casos são atribuídos aos argumentos (DPs) pelos núcleos lexicais de categoria V ou P e o núcleo funcional I (CHOMSKY, 1981, p. 49). O núcleo V atribui Caso Acusativo (ACC), o núcleo P, o Caso Oblíquo (OBL) e o núcleo I, o Caso Nominativo (NOM). Voltando ao exemplo (46), é possível observar que o predicador V (*comprar*) é um verbo transitivo que seleciona dois argumentos. Primeiramente, o núcleo V atribui ao seu argumento interno, *uma galinha*, o Caso ACC. Em português, esse Caso se relaciona à função sintática de objeto direto. O núcleo I atribui o Caso NOM ao argumento externo, *João*, que se moveu da posição Spec, VP, onde foi gerado, para a posição Spec, IP, para ali receber NOM sob Configuração Spec/Head (CHOMSKY, 1986, 1991).

Em algumas situações, o argumento externo permanece na posição em que foi gerado e, ali, recebe NOM do núcleo I por Regência, conforme proposta feita por Koopman e Sportiche (1991). Observe como ocorre essa atribuição na representação a seguir, do português europeu:

(47)



No português europeu, (cf. (47)), o argumento externo pode permanecer *in situ*, em Spec, VP (cf. COSTA, 1998). Nessa sentença, V se moveu até I, mas o sujeito não se moveu e, portanto, o Caso NOM foi atribuído por Regência.

Após a apresentação de algumas das teorias que representam módulos da UG, módulos esses que já nascem com o indivíduo e são, portanto, universais, na próxima subseção, farei algumas incursões no Parâmetro do Sujeito Nulo, tendo em mente como os sujeitos se realizam na posição sujeito, onde recebem Caso NOM.

2.2 O PARÂMETRO DO SUJEITO NULO

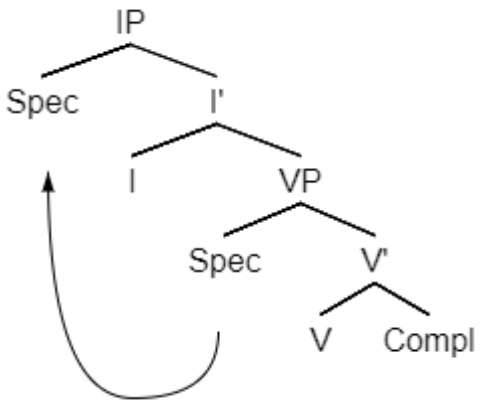
O Parâmetro do Sujeito Nulo (PSN) (do inglês, *Null Subject Parameter*) se relaciona diretamente com um princípio da GU, o EPP (do inglês *Extended Projection Principle*), também denominado de Princípio de Projeção Estendido. O Princípio de Projeção, conforme já apresentado, faz parte da Teoria X-barra, o qual estabelece que o núcleo lexical X terá a projeção intermediária X’ e a projeção máxima XP. Nessa hierarquização, o Caso NOM é sempre atribuído ao DP pelo núcleo funcional I, o que implica dizer que a posição de sujeito sempre é projetada, o que resulta no EPP proposto por Chomsky (1981):

(48) Princípio de Projeção Estendido (EPP)

A posição de sujeito é projetada em todos os níveis de representação

Essa posição de sujeito pode ser ocupada por argumentos externos (e internos, no caso dos verbos inacusativos) e por elementos que não são argumentos (os expletivos) e são inseridos na sintaxe diretamente nessa posição. Sobre os primeiros, o Caso NOM pode ser atribuído sob configuração Spec/Head, quando movido de sua posição de base (Spec, VP) para Spec, IP (cf. (49)) ou sob regência, quando esse movimento não ocorre, permanecendo o argumento externo em Spec, VP:

(49)



Em alguns casos, entretanto, o núcleo lexical V não seleciona argumentos, mas o Caso NOM continua sendo atribuído ao expletivo que ocupa essa posição. Isso acontece, por exemplo, em frases construídas com verbos existenciais ou que expressam fenômenos da natureza.

(50) a. It rained. (inglês)

b. Choveu. (português)

Em (50), o verbo *chover* não seleciona argumentos. No entanto, o expletivo na posição sujeito garante que o EPP seja satisfeito. No inglês, o expletivo é realizado foneticamente (*it*), ao passo que, em português, não o é e; portanto, é uma categoria vazia (*pro*).

A propriedade que o português e outras línguas têm de possuírem expletivo nulo está relacionada ao Parâmetro do Sujeito Nulo. Além disso, Rizzi (1982, 1986, 1988), reconhecendo que, no italiano, há sujeitos referenciais nulos, assumiu que línguas com a morfologia de flexão verbal (AGR, do inglês *Agreement "concordância"*) rica produzem esses sujeitos, de modo que o EPP é satisfeito pela categoria vazia *pro,* que recupera os traços número-pessoais do núcleo flexional I. Desse modo, o italiano e o PE são línguas *pro-drop* por permitirem sujeitos nulos, não-realizados foneticamente, e o francês e o inglês, que não possuem AGR rico, são línguas *não-pro-drop.* A seguir, é possível verificar o contraste entre o italiano (cf. (51a)) e o inglês (cf. (51b)):

(51) a. (Io) parlo italiano.
"Eu falo italiano."
b. I speak English.
"Eu falo inglês."

(RIZZI, 1986)

A realização fonética do sujeito pronominal de primeira pessoa (*Io*) é opcional no italiano; já, no inglês, é obrigatória.

Visando compreender as propriedades que compõem o PSN a partir da comparação entre diferentes línguas, Chomsky (1981) descreve um feixe de propriedades para as línguas *pro-drop*. Veríssimo (2017, p. 78-79), aplicando essas propriedades ao italiano e ao PE, observa o seguinte:

Propriedade 1: permitem a omissão do sujeito

(52) a. ___ parli italiano bravamente. [Italiano]
b. ___ falas italiano muito bem. [PE]

Propriedade 2: Permitem a inversão livre do sujeito em frase simples

(53) a. Ha telefonato Carlo a il suo padre [Italiano]
b. Telefonou o Carlos ao seu pai. [PE]

Propriedade 3: Permitem movimento longo de constituintes WH-

(54) a. Chi$_i$ pensi che ___$_i$ viene alla festa? [Italiano]
b. Quem$_i$ você acha que ___$_i$ vem à festa? [PE/PB]

Propriedade 4: Permitem pronomes resumitivos nulos em sentenças encaixadas

(55) a. Ecco la ragazza che mi domando che possa venire [Italiano]
b. Essa é a moça que acho que possa vir [PE/PB]

Propriedade 5: Permitem aparente violação do filtro *that-t* (“that-trace”)

(56) a. Chi$_i$ crede che ___$_i$ partirà? [Italiano]
b. Quem$_i$ você acha que ___$_i$ partirá? [PE/PB]

As propriedades acima não são verificadas no inglês, língua não-*pro-drop*, conforme verificado por Rizzi (1982), ao contrastá-la com o italiano, uma língua *pro-drop*:

a) O italiano permite o sujeito pronominal nulo seja referencial (cf. (57)) ou não-referencial (cf. (58)):

(57) a. ___ parla.
b. *___ speaks.
(58) a. ___ piove.
b. *___ is raining.

b) o italiano permite que objetos diretos em uma oração passiva permaneçam “*in situ*” (cf. (59)) e que argumentos externos de verbos permaneçam em Spec, VP (cf. (60)):

(59) a. ___ è stato dato *un premio* al presidente.

b. *___ was given *a prize* to the president.

(60) a. ___ ha telefonato *Gianni*.

b. *___ telephoned *John*.

c) o italiano permite o movimento do constituinte WH- para uma posição de sujeito mais alta, deixando a posição vazia na oração encaixada (cf. (61)):

(61) a. Chi credi che ___ telefonerà?

b. *Who do you think that ___ will telephone?

d) O italiano possui um AGR rico, pois os morfemas número-pessoas são capazes de identificar todas as pessoas do paradigma pronominal (cf. (62)):

(62) a. *It.:* parl-o, parl-i, parl-a, parl-iamo, parl-ate, parl-ano.

b. *Ing.:* speak, speak-s.

É importante dizer que os casos em que há sujeitos realizados foneticamente (os chamados sujeitos plenos) em línguas como o italiano estariam associados a contextos de ênfase, de contraste (GONÇALVES, 1994; RIZZI, 1988). Malta (2019, p. 78) observa:

> Conforme Rizzi (1988), em línguas como o italiano, a realização fonética do sujeito se torna redundante, visto que a desinência já se encarrega da distinção entre as pessoas verbais, o que ocasiona uma falta de necessidade do pronome. Por outro lado, sabe-se que no inglês a realização fonética do pronome é fundamental para que se realize a distinção entre as pessoas verbais, pois as desinências nesta língua não são suficientes para distingui-las, não havendo, assim, um acúmulo de informações desnecessárias. Desse modo, segundo o autor, numa língua como o italiano, o preenchimento do sujeito pronominal aconteceria somente para expressar ênfase ou contraste.

Tanto Chomsky (1981) como Rizzi (1982) defendem que os sujeitos nulos são licenciados por uma morfologia de flexão rica capaz de identificar as informações número-pessoais. Veríssimo (2017, p. 79) destaca, porém, que Chomsky, em 1981, observou que algumas línguas de sujeito nulo poderiam “não apresentar correlações absolutas a respeito dessas propriedades das línguas de sujeito nulo, mas sim uma

‘tendência’[11] para uma ou outra possibilidade do sistema”. Exemplos disso seriam o irlandês e o hebraico. De modo similar, Rizzi (1982, p. 144) já apontava para a necessidade de subparametrizar o PSN. Para tanto, ele fez, por exemplo, considerações sobre o valor pronominal das flexões, aproximando AGR da natureza dos pronomes clíticos, permitindo que a flexão fosse interpretada como argumento e tendo referencialidade. Mais recentemente, essa mesma proposta tem sido confirmada em trabalhos que analisaram mais línguas de morfologia de flexão rica (BARBOSA, 1995, 2009; ALEXIADOU; ANAGNOSTOPOULOU, 1998; PLATZACK, 2004; HOLMBERG, 2005; HOLMBERG; NAYUDU; SHEEHAN, 2009). A flexão verbal, uma vez tendo traços de pronome e referência, poderia permitir uma subparametrização, da seguinte forma:

> Traços da flexão:
> [+pronominal] e [+referencial] *pro-drop* em sentenças finitas.
> [-pronominal] e [-referencial] *não-pro-drop* em sentenças finitas.
> [+pronominal] e [-referencial] apenas *pro* expletivo.
>
> (RIZZI, 1982, p. 180)

Rizzi apresenta, então, que há línguas *pro-drop* que permitem sujeitos nulos referenciais e não referenciais, como o italiano, e aquelas que permitem apenas nulos não referenciais, o expletivo, como o alemão. Em 1986, Rizzi continua analisando a necessidade de identificar as condições que permitem *pro* nas línguas naturais. Ele, então, elenca as seguintes restrições:

a. *pro* deve ser licenciado. A existência de um “licenciador” é fruto de variação paramétrica entre as línguas, por exemplo, a flexão em I o núcleo lexical V podem reger *pro*;
b. *pro* deve ser identificado. Por esse motivo, *pro* sempre trará consigo os traços ϕ do elemento que o licencia. Por exemplo, os traços da flexão identificam *pro*.

[11] Destaque do autor da citação.

Toda essa discussão, no entanto, não abarcou a variedade de licenciamento de *pro* nas línguas naturais. Isso porque, mesmo definindo restrições para licenciar *pro*, Rizzi (1985, 1986) continuava as relacionando estritamente à morfologia de flexão verbal. Em 1984 e 1987, Huang, analisando o chinês, trouxe outra perspectiva que colaborou sobremaneira nos estudos sobre o parâmetro *pro-drop.* Essa língua não possui morfologia de flexão verbal e, nas frases, são presentes as raízes verbais[12]. Mesmo assim, sujeitos nulos são licenciados nessa língua. Os contextos, no entanto, são diferentes. Deus (2011, p. 229) comenta que Huang

> propõe a existência de um parâmetro tipológico que distinga as línguas naturais em dois tipos: (a) as línguas que são orientadas para o discurso (como o Chinês) e (b) as que são orientadas para a frase (como o Inglês).

Por exemplo, nas sentenças, a seguir, retiradas de Huang (1989, p. 188), o sujeito nulo da oração encaixada é controlado por um tópico já mencionado no discurso:

(63) Zhangsan shuo [***pro*/ta** lai le].
Zhansan disse [que ***pro*/ele** veio].

(64) Zhangsan xiangxin [***pro*/ta** hui lai].
Zhansan acredita [que ***pro*/ele** virá].

A respeito dessas duas sentenças, Huang (1989, p. 188) afirma:

> In both examples, the embedded subject may be null or non-null. If we assume that Case Filter applies to Chinese and that the subject of a clause is assigned Case only if it is governed by an element in INFL, then the possibility of having an embedded non-null subject shows that the subject position is governed, and the possibility of a null subject in the same position shows that Chinese is a pro-drop language.[13]

[12] No chinês, seja escrito ou em suas muitas variedades faladas, não há nenhum sistema desinencial que confira ao componente verbal conceitos de tempo, modo, número ou pessoa. Essas informações, quando relevantes à comunicação, serão apresentadas por meio de partículas, de adjuntos ou do contexto, ou seja, serão propiciadas pelo discurso.

[13] "Em ambos os exemplos, o sujeito incorporado pode ser nulo ou não. Se assumirmos que o Filtro de Caso se aplica ao chinês e que o sujeito de uma sentença recebe Caso apenas se for regido por um elemento do domínio flexional I, então a possibilidade de haver um sujeito não nulo incorporado mostra que a posição de sujeito é regida, e a possibilidade de um sujeito nulo na mesma posição mostra que o chinês é uma língua *pro-drop*." Tradução minha.

Desde a proposta de Huang, vários pesquisadores, inclusive Rizzi, revisaram e expandiram suas investigações sobre o PSN. Segundo Deus (2009, p. 232), esse parâmetro foi revisitado, reinterpretado, abandonando-se “a ideia convencional de que todas as línguas são “bem-comportadas” no que toca ao PSN”.

A partir dos anos 2000, estudos importantes foram realizados, analisando uma quantidade considerável de línguas diversas (KATO, 2000; NEWMEYER, 2004; NICOLIS, 2008). Essas novas perspectivas permitiram que Holmberg e Roberts (2009) propusessem quatro sistemas de sujeito nulo, reunindo nelas diferentes feixes caracterizadores do PSN. Verissimo (2017, p. 83) retoma essa proposta e resume essas propriedades:

- *Línguas canonicamente pro-drop*: há uma morfologia verbal rica suficiente para possibilitar a ampla ocorrência de sujeitos nulos em todas as pessoas gramaticais e em contextos variados (ex.: italiano, português europeu, grego e turco)
- *Línguas parcialmente pro-drop:* os sujeitos sem referência definida tendem a ser nulos, mas há fortes restrições à ocorrência de sujeitos nulos referenciais (ex.: hebraico, russo, finlandês e português brasileiro);

- *Línguas radicalmente pro-drop:* línguas que permitem sujeitos nulos sob condições de natureza discursiva e não apresentam qualquer tipo de marca verbo-flexional indicativa de concordância (chinês, japonês, tailandês e vietnamita);

- *Línguas de expletivos nulos:* não há a possibilidade de sujeitos nulos referenciais, mas de expletivos nulos (ex.: alemão, holandês e o crioulo de Cabo Verde).

Veríssimo afirma, ainda, que permanece a binaridade do PSN, tomando por base a proposta de Holmberg e Roberts (2009), via hierarquia paramétrica, que será percorrida pela criança durante a aquisição para efetivar a marcação do valor paramétrico correspondente à sua língua. Retomando a proposta de Holmberg e

Roberts (2009), Veríssimo (2017, p. 84) esquematiza a seguinte hierarquia para o PSN:

Há sentenças com sujeitos nulos em algum tipo de contexto?
NÃO
Língua *não-pro-drop*
Inglês
Francês

SIM
Há flexão verbal?
NÃO
Língua radicalmente *pro-drop*
Chinês
Japonês

SIM
Há expletivos nulos ocorrendo livremente?
NÃO
Língua de expletivo nulo
Alemão
Caboverdiano

SIM
Há sujeitos referenciais nulos ocorrendo livremente?
NÃO
Língua *pro-drop* parcial
Russo
Português brasileiro
Hebraico

SIM
Língua *pro-drop* canônica
Italiano
Português europeu
Espanhol

Centrando minha atenção agora no PB e no PE, verifico que, antes da proposta de Holmberg e Roberts (2009), Duarte (1995), ao perceber a diferença entre essas variedades do português no que se refere ao comportamento dos sujeitos nulos, associou esse comportamento ao paradigma de flexão verbal: no PB a morfologia vem sendo enfraquecida, ao contrário do que se verifica no PE (cf. também GALVES, 2001). Analisando o pouco percentual de sujeitos nulos em sua pesquisa, realizada na segunda metade do século XX, sendo que os nulos compreenderam apenas 29% do total de sujeitos que analisou, no PB, Duarte propôs que essa língua apresentava um sistema de sujeito nulo agonizante, por estar atravessando um processo de mudança de língua *pro-drop* para *não-pro-drop* e, portanto, “segue uma luta em direção ao sujeito pleno” (DUARTE, 1995, p. 152). Galves (1998, 2001) menciona, por exemplo, que o verbo de terceira pessoa, no PB, não tem sido suficiente para uma interpretação

definida do sujeito, de modo que os brasileiros têm realizado esses sujeitos por meio de pronomes ou sintagmas nominais.

No entanto, estudos mais recentes que analisam a fala e a escrita do PB, diacrônica e sincronicamente, têm revelado que não existe uma tendência para a mudança supracitada (MODESTO, 2004; KAISER, 2006; OLIVEIRA JUNIOR, 2018), pois os sujeitos nulos encontram-se submetidos a restrições observadas em outras línguas de sujeito nulo, conforme pontuam Kato e Duarte (2014). Segundo essas autoras, “o Português Brasileiro (PB) vem mostrando um declínio na ocorrência do sujeito nulo referencial definido ao contrário do que revela o Português Europeu (PE)” (KATO; DUARTE, 2014, p. 2), conforme evidenciado a seguir:

(65) a. [Minha esposa]$_i$ trabalha na Embratel. Ela$_i$ ganha bem, mas eu acho que ela$_i$ devia ganhar mais porque ela$_i$ merece. (PB)

b. Ele$_i$ quer pescar tudo; Ø$_i$ quer sempre arranjar umas taças. E Ø$_i$ tem tido sorte com isso, porque Ø já teve três (taças) e eu inda só tive uma, que foi nesse concurso. (PE)

Em (65a), temos a realização fonética do sujeito referencial em todas as orações, mesmo sendo correferentes; já, no PE, (65b), após a primeira menção do referente que ocupa a posição sujeito e se realiza foneticamente na primeira oração, a posição sujeito das orações subsequentes é ocupada pelo sujeito nulo que mantém correferência com o sujeito da primeira oração.

Holmberg, Nayadu e Sheehan (2009) e Holmberg (2010) verificam que sujeitos nulos de referência indefinida/arbitrária são encontrados no PB em contextos não verificados no PE:

(66) a. É assim que *pro* faz o doce. (PB)

b. É assim que *se* faz o doce. (PE)

(HOLMBERG, 2010, p. 92)

Como em (66), casos de sujeitos pronominais indefinidos, que são nulos no PE, podem não ser nulos no PB. A respeito disso, Krzek (2017, p. 308) verifica que:

> In consistent NSL all subject pronouns regardless of the person and tense can be null. Indefinite null subjects, on the other hand, have to be overt. Holmberg (2010a: 92) illustrates this difference by contrasting Brazilian Portuguese (BP), a partial NSL, with European Portuguese (EP), a consistent NSL.[14]

Segundo Kato e Duarte (2014), o uso de sujeitos nulos com referência arbitrária/indefinida pode ser explicado pela hierarquia de referencialidade proposta por Cyrino, Duarte e Kato (2000). Sob a perspectiva dessa hierarquia,

> Argumentos [+N, +humano] estão no extremo mais alto na hierarquia referencial, enquanto não-argumentos estão na posição mais baixa. Com relação aos pronomes, o falante (*eu*) e o interlocutor (*você*), sendo inerentemente argumentos humanos, primeira e segunda pessoas pronominais, estão no ponto mais alto na hierarquia; a terceira pessoa se situa num ponto mais baixo, devido à interação de traços [+/-humano] e [+/específico]. O sujeito que se refere a uma proposição (o sujeito neutro) está numa posição ainda mais baixa. (KATO; DUARTE, 2014, p. 4-5)

A proposta feita a partir dessa hierarquia é que, quanto mais referencial o pronome fosse, maior a probabilidade de ele ser realizado. Dessa forma, aplicando a hierarquia da referencialidade aos dados do PB, as autoras observaram que ainda há pronomes nulos sem restrição.

Assim, as autoras chegam à seguinte conclusão: “o PB é pautado não pelo princípio ‘Evite Pronome’, mas pelo princípio parcial ‘Evite pronomes referencialmente deficientes’” ou não-referenciais (KATO; DUARTE, 2014, p. 7).

Ademais, Kato e Duarte (2014) verificam que há contextos com expletivo nulo no PB (cf. (67)) que se alterna com construções pessoais em que a posição sujeito é ocupada por um constituinte que se move para essa posição (cf. (67b)):

(67) a. *pro* chove muito nessas florestas.
b. Essas florestas chovem muito

Um fato curioso levantado pelas autoras é que o exemplo (67b) foi extraído de uma redação de vestibular, língua escrita monitorada, o que evidencia uma tendência ao preenchimento da posição Spec, IP por sujeitos plenos.

---

[14] “Em línguas de sujeito nulo consistente, todos os sujeitos pronominais, independentemente da pessoa e do tempo verbal, podem ser nulos. Sujeitos nulos indefinidos, por outro lado, precisam ser plenos. Holmberg (2010a: 92) ilustra essa diferença contrastando o português brasileiro (PB), uma língua de sujeito nulo parcial, com o português europeu (PE), uma língua de sujeito nulo consistente.” – tradução minha.

Outra restrição apresentada por Kato e Duarte (2014) para a produção de sujeitos nulos no PB é o contexto de orações encaixadas em que os referentes desses sujeitos podem ser retomados na oração principal (cf. (68a-b)) ou no discurso (cf. (68c-d)):

(68) a. João$_i$ disse que ele$_i$ vai à festa.
b. João$_i$ disse que *pro*$_i$ vai à festa.
c. João$_i$ disse que ele$_j$ vai à festa.
d. João$_i$ disse que *pro*$_j$ vai à festa.

Sendo assim, concluem Kato e Duarte (2014, p. 17) que

> A sintaxe do PB vem mudando e que a variação exibida não é apenas efeito de uma mudança em curso, mas que ela pode ser explicada como propriedades de uma gramática estável caracterizável em termos de restrições semânticas de referencialidade e de propriedades morfossintáticas. (KATO; DUARTE, 2014 p. 17)

Essa conclusão, além de corroborar a proposta inicial de Holmberg e Roberts (2009), aplicada ao PB por Holmberg, Nayadu e Sheehan (2009), representa uma definição no sentido contrário ao que Duarte propôs em 1995 e mostra que o PB é uma língua de sujeito nulo parcial.

Após essa abordagem sobre o PSN, considero pertinente distinguir aquisição de aprendizagem de uma língua, com o objetivo de situar a realidade linguística do português em Timor-Leste.

2.3 A AQUISIÇÃO DA LINGUAGEM

Do ponto de vista da teoria gerativa, especificamente do modelo P&P, a aquisição de uma língua natural ocorre via fixação dos valores paramétricos na GU do falante. Ela acontece a partir da exposição da criança ao *input*. Claramente, esse processo é chamado de aquisição e não de aprendizagem, conforme veremos adiante, porque envolve uma estrutura existente na mente-cérebro que promove a aquisição de um conhecimento linguístico internalizado sobre uma ou mais línguas que independe de instrução formal. Segundo Pollock (1997), a aquisição de sua própria

língua materna não é algo que as crianças façam, mas que lhes acontece. Ou seja, não adquirimos uma língua zero. Nosso aparato mental, a Faculdade da Linguagem, já é preparada para desenvolver essa habilidade, tal qual não nascemos enxergando, mas o nosso sistema sensorial já está preparado para desenvolver, também, essa habilidade. Para Kail (2013, p. 11-12),

> A maioria das crianças domina as estruturas básicas de sua língua materna por volta dos 4 anos, ao mesmo tempo em que dá provas de desempenhos expressivos em outros campos do desenvolvimento cognitivo e social, que parecem exercer um papel de destaque na emergência da própria linguagem. Como resultado desse processo prodigioso, a criança de 4 anos é um ser extremamente sofisticado. Como é que tamanha façanha pode ser realizada normalmente, em um prazo tão curto, por todas as crianças e em todas as culturas? É difícil evitar a conclusão de que a linguagem é parte de nossa herança biológica, uma realização que depende das características específicas do cérebro humano.

Essa regularidade mencionada por Kail é uma das evidências para o entendimento de que a aquisição é um processo uniforme para todas as crianças, salvo os casos de sérias patologias ou limitações motoras ou sensoriais. Além disso, para que ocorra essa aquisição, essas crianças precisam somente estar expostas aos dados linguísticos primários (o *input*), não passando por treinamento ou ensinamento explícito do conhecimento linguístico que está adquirindo.

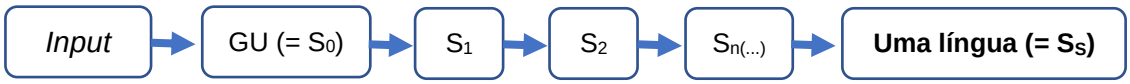
O que chama atenção a respeito do *input* é que é limitado. Ou seja, a criança nunca será exposta à língua em sua totalidade, mas adquirirá um conhecimento que vai muito além da informação disponível no ambiente. Ela alcança o todo, sendo exposta às partes e pode produzir estruturas nunca antes produzidas pelos adultos à sua volta, evidenciando assim a criatividade linguística, característica da linguagem humana.

Chomsky (1971) argumenta que o nosso uso da linguagem é criativo, ou gerativo: a partir de um conjunto infinito de morfemas ou de palavras, podemos criar um conjunto infinito de frases jamais ouvidas antes. Ao mesmo tempo, o uso da linguagem é econômico: podemos encaixar as frases umas nas outras, utilizando-nos da propriedade da recursividade, que também é própria da linguagem humana.

Um argumento muito forte na direção da aquisição como expressão do inatismo linguístico é o da pobreza de estímulo. “O *input* fornecido à criança pelo ambiente linguístico seria muito incompleto e inconsistente para determinar acesso à complexidade gramatical” (KAIL, 2013, p. 52). Os estudos mostram que há um período

crítico de aquisição que vai desde os primeiros contatos da criança com a língua do ambiente até cerca dos seus 4 ou 5 anos de idade (MOGFORD; BISHOP, 2002). Nesse interstício cognitivo de *input* finito a *output* infinito, está a infinitude discreta da linguagem humana só conseguida a partir da GU que filtra do *input* o que é necessário, processa e constrói um conhecimento linguístico complexo e sofisticado. Para tanto, é necessário que, nesse período crítico, haja, conforme já enunciado, a fixação dos valores paramétricos que irão definir as diferentes gramáticas e é exatamente essa valoração que a criança, no período crítico da aquisição irá receber a partir do *input* da língua-E. Vejamos a figura, a seguir, que esquematiza o caminho do estágio inicial ao estágio estável:

Figura 2 - Processo de aquisição no período crítico



Fonte: autor deste trabalho.

Conforme a figura 2, a criança parte do estágio inicial (S_0) e fixa todos os parâmetros até atingir o estágio estável (S_S) (do inglês, *stable stage*). Nesse momento, a criança apresenta uma gramática nuclear (*core grammar*) próxima à do falante adulto. Ela fará isso, da mesma maneira, para uma segunda ou terceira língua, em casos de contexto bilíngue ou multilíngue.

No processo de aquisição, a criança parte, inicialmente, do reconhecimento dos sons de sua língua. Ela os percebe e depois consegue segmentá-los a partir do léxico que está sendo aprendido. As palavras, por sua vez, à medida que vão sendo aprendidas com suas propriedades léxico-semântico-funcionais, vão permitindo à mente da criança sua estruturação hierárquica, sua sintaxe. Primeiramente, as regularidades são fixadas e, em seguida, irregularidades e sintagmas de maior complexidade. Com o decorrer do tempo, vão surgindo frases cada vez mais complexas a ponto de permitir à criança a construção de enunciados elaborados, utilizando, além da criatividade e da recursividade, os procedimentos de referenciação e anáfora (MIOTO, 1995).

O entendimento do processo de aquisição se revela importante porque tem relação direta com a mudança linguística. Segundo Lightfoot (1991), a mudança linguística se inicia com a criança. Se um novo valor é fixado por ela, sendo que ele não está previsto na gramática adulta, uma nova gramática começa a emergir. Muitos aquisicionistas defendem essa ideia, sobretudo, quando o *input* a que a criança está exposta é variável. Embora a mudança superficial se evidencie nos contextos de uso de falantes adultos, é, portanto, o *input* variável, chegando às próximas gerações pela nova fixação paramétrica nos estágios da aquisição, que determinará mudanças efetivas na gramática das línguas (GALVES, 2012).

2.3.1 Aquisição *versus* aprendizagem

Conforme apresentei na subseção 2.2, a aquisição de uma (ou mais de uma) língua natural (também considerada língua materna (LM)) é o processo natural pelo qual passa toda criança durante o período crítico em que os valores paramétricos são fixados (CHOMSKY, 1981; MADEIRA, 2004). Diante disso, opto por fazer uma distinção entre uma LM e uma língua não-materna (LNM), tomando por base Almeida (2008).

Diferentemente da LM, a LNM não é adquirida, mas é resultante de outro processo, o de aprendizagem. Conforme citado por Almeida (2008 p. 46), “o termo *Aquisição* está reservado para a LM; todas as restantes línguas *aprendem-se*, não se *adquirem*” [grifos do autor]. Almeida (2008) ainda afirma que

> o termo *Aprendizagem* implica o estudo formal da língua, geralmente em aula, com o professor, com explicações gramaticais, exercícios e simulações de situações para fomentar o diálogo (p. 46).

Nesse sentido, para os casos em que uma língua é apresentada à criança, ao jovem ou adulto, fora do período crítico de aquisição ou em contextos limitados de *input*, que não o familiar, por exemplo, haverá a aprendizagem de uma LNM, um processo ativo de construção de conhecimento de outra língua além da materna (MADEIRA, 2004).

Vale ressaltar que a LM que menciono aqui não é necessariamente *uma* língua. A criança, no período crítico de aprendizagem, pode estar em contato intenso com o

input de mais de uma língua e, portanto, pode adquirir tantas quantas elas forem. Isso não é incomum em contextos de bilinguismo ou em que o contexto familiar tem línguas diferentes sendo usadas em concorrência no dia a dia. De acordo com Pupp Spinassé (2006, p. 5), a LM tampouco se resume a

> apenas uma língua. Normalmente é a língua que aprendemos primeiro e em casa, através dos pais, e também é frequentemente a língua da comunidade. Entretanto, muitos outros aspectos linguísticos e não-linguísticos estão ligados à definição. A língua dos pais pode não ser a língua da comunidade, e, ao aprender as duas, o indivíduo passa a ter mais de uma L1 (caso de bilinguismo). (...) A título de ilustração: uma criança nasce e cresce na Alemanha, filha de um francês com uma colombiana. Se com cada um dos pais ela se comunica nas suas línguas respectivas, e na creche, na rua, com os amigos e vizinhos o alemão é a língua diária, essa criança tem, claramente, três línguas maternas: francês, espanhol e alemão. (PUPP SPINASSÉ, 2006, p. 5)

Por outro lado, Pupp Spinassé (2006) não diferencia aquisição de aprendizagem. Ela, assim como outros autores, considera o processo de apropriação de uma segunda língua como aquisição. Alguns desses autores também consideram que é possível adquirir uma segunda língua tardiamente de maneira natural, similar à maneira como se adquire uma LM. (SLABAKOVA, 2016; MARCELINO, 2017).

A título de fixar uma filiação teórica, no entanto, ressalto que, ainda que se possa afirmar que uma LNM *possa* ser adquirida, por um processo similar à aquisição de LM, acredito haver necessidade de diferenciação quanto ao *status* dessa língua no contexto do falante. O caso que aqui chamo atenção é que, *se* essa segunda língua, que não é a materna, tem utilização em contextos limitados, e seu *input* não é robusto, ela tampouco pode ser adquirida de maneira natural.

Uma das evidências apontadas no estudo de Marcelino (2017, p. 50) para a aquisição natural de uma L2 inclui o surgimento de propriedades não aprendidas formalmente. Ele diz: “o aparecimento de propriedades abstratas, sobre as quais o aprendiz não recebeu instrução formal, pode ser um indicativo de que há princípios da GU norteando ou restringindo as possíveis gramáticas de L2.” Quanto a esse argumento, concordo plenamente. Como sustentarei mais à frente, para a aprendizagem de uma LNM, há a necessidade de interferência da GU, como o caminho para se chegar à aprendizagem efetiva. A essa altura, portanto, é apropriado que sejam vistas as características desse grupo de línguas a que me refiro como LNM.

As LNM se diferenciam de vários modos (ALMEIDA, 2008). Podem ser, por exemplo: a) línguas adicionais em contextos plurilíngues, aqueles em que várias línguas são utilizadas em várias regiões de certo território, mas que, no seio familiar, o uso é de uma língua apenas, a LM; b) línguas com as quais a criança entrará em contato apenas por meio de um processo de instrução formal, como o ambiente escolar ou qualquer outro que ofereça um ensino sistemático da língua; e c) línguas estrangeiras aprendidas, não pertencentes, inicialmente, ao contexto de aquisição e de uso natural da linguagem pela criança. Nesses casos, aplicam-se os conceitos de L2, L3 e assim por diante a qualquer LNM, representando esse número a ordem em que foi aprendida. Vale referir que uma L2 pode ser LNM do próprio contexto geográfico em que vive o indivíduo ou, de fato, uma LE, que não pertence a esse contexto.

Desse modo, Almeida (2008, p. 48) apresenta distinções entre o desenvolvimento de uma LM (L1) e de uma LNM (L2), conforme quadro 9, a seguir:

Quadro 9 - Desenvolvimento de L1 e L2

Desenvolvimento da L1	Desenvolvimento da L2
Salvo raras excepções, o sucesso é garantido.	Poucas probabilidades de sucesso completo na aprendizagem, no caso da L2, com ocorrência frequente de fossilização, ou seja, de persistência no uso de estruturas desviantes, mesmo em níveis muito avançados de proficiência.
_____	A língua materna do aprendente desempenha um papel na aprendizagem.
Maior tendência para a uniformidade.	Caracteriza-se pela variabilidade, quer entre aprendentes, quer em aprendentes individuais.
Não é influenciado por factores individuais.	É influenciado por factores individuais – motivação, atitudes, aptidão linguística, etc.
Não é aparentemente influenciado pelo ensino formal e pelas correcções.	É provavelmente influenciado pelo ensino formal e pelas correcções.

Fonte: Almeida (2008, p. 48)

De acordo com o quadro 9, a influência da LM na aprendizagem de uma nova língua é presente. Essa influência, natural, faz com que características da LM sejam refletidas, ainda que temporariamente, na língua-alvo da aprendizagem. O processo chamado, por alguns pesquisadores de Aquisição de Segunda Língua (ASL) – aqui aquisição significando, de fato, aprendizagem – de interlíngua, é a competição entre a gramática internalizada do falante em contato com outra que lhe é externa, ocorrendo transferência de características daquela para esta. Ao longo da aprendizagem, no entanto, essas influências, que caracterizam erros de aprendizagem, tendem a ser corrigidas ou amenizadas por meio das correções próprias do ensino formal. Essa mesma perspectiva é defendida por Magalhães (2000)[15], ao esclarecer que

> a criança que já passou pelo processo de aquisição vem para a escola com um conhecimento de língua nativa (Língua-I) pronto e, muitas vezes, ao chegar à escola, é apresentada a formas que não correspondem àquelas que ela adquiriu. Mesmo diante de formas diferentes, a criança vai utilizar o conhecimento que já dispõe e a escola vai tentar reprimir esse uso através de correções. (MAGALHÃES, 2000, p. 43)

Por outro lado, em algumas situações, os erros de aprendizagem se solidificam ou são fossilizados. Por fossilização "entende-se um mecanismo através do qual o aprendente tende a conservar na sua interlíngua certos itens, regras e subsistemas linguísticos da sua língua materna" (OTTONELLO, 2005, p. 378 *apud* ALMEIDA, 2008, p. 49). O fenômeno da fossilização é comum, apresentando, em línguas aprendidas, algumas características da LM dos falantes.

Segundo Freire (2013, p. 44), o período em que ocorre a aprendizagem também pode significar maior sucesso para o indivíduo que desenvolve uma LNM:

> A ideia de que pessoas mais jovens têm melhores condições para adquirir uma segunda língua é resultado da teoria de "período crítico", proposta por Lenneberg (1967); Halle (1962); King (1969); Wilkins (1972), entre outros, os quais argumentam que a aquisição da segunda língua por adultos é inferior, do ponto de vista qualitativo, quando comparada com a aquisição de língua materna. Esse ponto de vista se embasa no pressuposto de que o indivíduo é programado biologicamente para adquirir a língua antes da puberdade. Assim, o adulto não pode recorrer a capacidades inatas de aquisição.

[15] Em sua pesquisa, Magalhães (2000) sustentou que o sujeito nulo, no PB é ensinado na escola, de modo que a instituição escolar tenta reverter a inovação da criança em direção ao sujeito pronominal pleno, para que ela produza mais sujeitos nulos.

Assim, a ideia que defende Freire, e à qual me filio, é que, quanto mais distante do período crítico de aquisição for a aprendizagem, mais dificultosa ela se revelará, assim como as possibilidades de fossilização aumentarão.

As possibilidades de sucesso na aprendizagem podem estar relacionadas ao papel desempenhado pela GU nesse processo, que é diferente da aquisição. Conforme Freire (2013) apresenta, Finger (2003) elencou três hipóteses para o acesso à GU na aprendizagem de LNM. São elas:

a) A Hipótese de Acesso Nulo (HAN)

Nesta hipótese, a aquisição de uma LM e de uma LNM são vistas como processos cognitivos distintos. Desse modo, o adulto não tem acesso à sua GU durante a aprendizagem.

b) A Hipótese de Acesso Parcial (HAP)

Nesta hipótese, há a ação parcial da GU na aprendizagem de LNM. Esse acesso à GU se dá, por sua vez, por intermédio da LM, com valores parâmetros já estabelecidos, que servirão de base para a aprendizagem da outra língua. Assim, a LM seria a base para o desenvolvimento da LNM, valendo referir que, se a LNM possuir configurações paramétricas diferentes, outros mecanismos serão necessários para auxiliar na aprendizagem e na geração interna das sentenças.

c) A Hipótese de Acesso Total (HAT)

Nesta hipótese, não há diferenças entre a aquisição e a aprendizagem no que diz respeito ao acesso à GU e descarta-se a existência de um período crítico. Segundo essa perspectiva, no período de aprendizagem, haveria continuidade no acesso dos princípios e parâmetros disponíveis na GU.

Em sua pesquisa sobre a perda de características de sujeito nulo da LM, no caso o espanhol, por conta de o falante viver em contexto de uma LNM, o PB, Freire (2013) adota a HAT para desenvolver o seu trabalho. No entanto, nesta dissertação, ao analisar dados produzidos por estudantes universitários timorenses, afasto-me

dessa hipótese, pois, adotando aqui que aquisição e aprendizagem são processos diferentes, assumo a existência do período crítico de aquisição. Filio-me, portanto, à HAP, considerando que, uma vez passado o período crítico, qualquer aprendizagem de língua acontecerá com o acesso à GU após o estabelecimento do seu estágio estável, alcançado na aquisição da LM, e no qual os valores paramétricos encontram-se fixados, conforme defendem Schachter (1989) e Strozer (1992).

2.3.1.1 O caso específico de Timor-Leste

A realidade linguística de Timor-Leste se insere de maneira muito profícua quando se diferenciam a aquisição e a aprendizagem. No território de Timor, são faladas várias línguas maternas, cerca de dezesseis (ALBUQUERQUE, 2011). Costa (2005, p. 614-615) verifica que:

> O estrangeiro que queira percorrer o território e tentar comunicar com os seus habitantes vê-se perante uma babélica imagem de Timor e tem de recorrer a um intérprete, já que não pode estabelecer um verdadeiro entendimento entre indivíduos de grupos diferentes, cada um falando a sua língua materna. Assim, há que recorrer ao tétum, língua comum à maioria dos grupos. As línguas de Timor, umas apresentam afinidades com a família austronésia, tais como o Ba'ikeno, Manbae, Galolen, Tetun, etc., sendolhe outras estranhas, geralmente classificadas como papuas, tais como bunak, fatluku, makasae, etc.

Como dito por Costa (2005), o tétum é a língua comum, além de ser a LM de parte considerável da população. Para além disso, também é a Língua Oficial (LO), *status* que compartilha com o português. Almeida (2008, p. 33) menciona que, de acordo com a Unesco, LO é:

> A língua utilizada no quadro das diversas actividades oficiais: legislativas, executivas e judiciais de um Estado soberano. É qualquer língua que um Estado defina como tal, quer na Constituição, quer por lei ordinária, quer apenas por via do costume.

No que se refere ao tétum e ao português, no entanto, há uma importante diferenciação: o tétum é a língua nacional de Timor-Leste, muitas vezes usada como língua franca. É também a LM de certa parcela da população do país, especialmente na região da capital, Díli. O português, por questões históricas e sociais, tem um uso

muito mais restrito. É a língua do Estado, da administração pública, dos contratos, dos atos normativos, dos discursos oficiais e da escolarização. Ou seja, está nos contextos formais, praticamente fora do âmbito familiar. Sendo assim, o português em Timor-Leste é uma LNM e seu contexto de ensino formal, em alguns casos, é semelhante ao de uma língua estrangeira, pois o primeiro contato que alguns têm com a língua é na escola. Por conta dessa situação, acontecem, aos timorenses, dificuldades relativas à aprendizagem do português. Isso é demonstrado no tocante à aprendizagem do paradigma de flexão verbal ensinado a partir da norma europeia do português:

> Um problema frequente que é comum a praticamente todos os aprendentes da língua portuguesa em Timor-Leste, que dá origem a muitos dos erros na interlíngua e que se apresenta a título exemplificativo é, como identifica Esperança (2001), a complexidade das conjugações verbais. Os verbos em tétum e bahasa não têm variação de pessoa, número, modo, aspecto. No tétum, o tempo é marcado por palavras próprias como *ona* (passado), *sei* (futuro), *atu* (na eminência de ser feito ou acontecer)... Os tempos compostos do português levantam também problemas. (ALMEIDA, 2008, p. 49)

Outros problemas que diferenciam o tétum do português, de acordo com Batoreo (2010), são a ausência do verbo copulativo *ser*, a ausência do artigo definido, a colocação do artigo indefinido posposto ao nome (ex. “vou para casa *uma*” em oposição ao que ocorre no português “vou para *uma* casa”) e a formação de expressões como respostas a perguntas na negativa (ex. “— Não vai sair mais tarde? — Sim”, confirmando a negativa em oposição ao que ocorre em português “— Não vai sair mais tarde? — Não”). Essas diferenças constituem desafios para o ensino do português em Timor-Leste.

Diante dessa perspectiva, minha hipótese é a de que, no caso do PSN, também haveria a dificuldade do estudante timorense. Isso, por sua vez, leva-me ao seguinte questionamento: convivendo num contexto multilíngue, no qual estão línguas como o tétum, língua não-*pro-drop*, como se daria a aprendizagem do sistema *pro-drop* do PE por um timorense, resultando no PTL?

3 PROCEDIMENTOS METODOLÓGICOS

Neste capítulo, apresentarei os procedimentos metodológicos a que submeti a pesquisa, incluindo, desde os métodos utilizados, a coleta dos dados e a constituição do *corpus*, especialmente no que tange aos critérios de inclusão e exclusão, até a seleção das variáveis e ao tipo de análise dos dados.

Lakatos e Marconi defendem (2003, p. 83) que “não há ciência sem o emprego de métodos científicos”. Essa assertiva invoca a responsabilidade com a qual os trâmites procedimentais da pesquisa devem ser levados a desenvolvimento, conferindo a ela autenticidade, confiabilidade e credibilidade.

3.1 TIPO DE MÉTODO

Para este estudo, empreguei como método de abordagem o hipotético-dedutivo que, de acordo com Bunge (1974), parte da colocação de um problema que é analisado de acordo com determinado modelo teórico. São deduzidas, de acordo com esse modelo, as consequências particulares, extraindo de suportes empíricos dados para o teste das hipóteses. Os resultados são obtidos, dessa forma, a partir da interpretação dos dados à luz do modelo teórico. Adotando para a análise a Teoria Gerativa, em específico, o Modelo de Princípios e Parâmetros (CHOMSKY, 1981, 1986), analisei a concordância e o PSN em dados escritos do Português de Timor-Leste (PTL), assumindo, portanto, a natureza modular da Faculdade da Linguagem e a existência da GU já presentes na mente-cérebro do falante, que são capazes de explicar os mecanismos internos ativados na geração das sentenças nessa variedade do português, e, em particular, de como a posição sujeito é preenchida, tendo como hipótese a possível influência do contexto multilíngue de Timor-Leste no PTL.

Tendo em vista que o português não é LM dos timorenses e que seu aprendizado se assemelha ao de uma LE, a hipótese é que os dados do PTL resultam de aprendizagem, e não de aquisição. Ao mesmo tempo, analisar a natureza de concordância e como ocorre o Sujeito Nulo nessa variedade do português permitirá, ainda, compreender como funciona a Hipótese do Acesso Parcial à GU, que toma

como base o conhecimento linguístico já adquirido para a aprendizagem de uma LNM (SCHACHTER, 1989; STROZER, 1992).

Como método de procedimento, adotei o método estatístico, que permite “obter, de conjuntos complexos, representações simples e constatar se essas verificações simplificadas têm relações entre si” (LAKATOS; MARCONI, 2003, p. 108). Desse procedimento estatístico, objetivo extrair os quantitativos gerais do tratamento dos dados, bem como os valores percentuais das ocorrências do fenômeno linguístico em análise no *corpus*. Assim, é possível descrever, analisar e interpretar quantitativamente os dados.

3.2 COLETA DE DADOS

Os dados coletados foram extraídos de redações produzidas no ano de 2018 por estudantes timorenses que se submeteram a um exame para ingresso na Universidade da Integração Internacional da Lusofonia Afro-brasileira (UNILAB), localizada em Redenção-CE. Essas redações encontram-se arquivadas em um setor próprio dessa universidade, configurando-se, portanto, esta pesquisa, além de quantitativa, como documental, pois os dados submetidos à análise decorrem de fontes primárias que evidenciam um registro linguístico produzido em um contexto de maior monitoramento da língua.

Vale referir ainda que a UNILAB, implantada em 2010, é uma instituição pública criada para promover a integração internacional entre os países onde se fala o português. Esses países, membros da CPLP, dentre os quais figura Timor-Leste, enviam estudantes para cursar, no Brasil, cursos de áreas de interesse na cooperação internacional. No caso específico de Timor-Leste, essa parceria significa capacitação de pessoal para atuar nas mais diversas funções, tendo em vista a carência que aquele país ainda tem no que se refere à formação profissional de nível superior.

A UNILAB surge, então, como um importante agente político e social na integração dos Países Africanos de Língua Oficial Portuguesa (PALOP) e do Timor-Leste com o Brasil. De forma emblemática, essa universidade tem sua sede no Ceará, na cidade de Redenção, o primeiro município brasileiro a abolir a escravidão oficialmente. Além dos seus *campi* no Ceará (nas cidades de Redenção e Acarape), a Universidade conta com um *campus* na cidade de São Francisco do Conde, na

Bahia. Segundo Malta (2019, p. 108), os cursos ministrados nessa instituição têm ênfase em “formação de professores, desenvolvimento agrário, processos de gestão e saúde pública, engenharia”, entre outros.

Como requisito para ingresso na UNILAB, o processo seletivo inclui a escrita de uma redação em português pelos candidatos ainda em seu país de origem. Essa seleção ocorre de maneira presencial e as provas são enviadas ao Brasil para a avaliação. Após os procedimentos avaliativos, as redações ficam armazenadas na instituição e podem ser utilizadas para fins culturais e científicos. O objetivo dessa prova é avaliar a expressão escrita do candidato na elaboração de um texto dissertativo diante de um tema proposto.

Os critérios de correção, de acordo com a UNILAB, envolvem a adequação ao tema, a sequência argumentativa, a norma padrão da língua portuguesa e os requisitos para a coesão e a coerência textuais. Além disso, no ano da seleção em que foram produzidas as redações cujos dados estão submetidos à análise nesta pesquisa, a proposta de redação esteve voltada à temática: “A língua portuguesa e o Timor-Leste”. Em vista disso, foram recorrentes no *corpus* frases com considerações dos escreventes sobre a relevância do português para si e/ou para os timorenses, menções ao contexto multilíngue, de políticas linguísticas e até mesmo à relação pessoal do escrevente com os idiomas que fala.

Para que tivéssemos acesso às redações dos estudantes timorenses, entrei em contato com a mestra Maria Joseane Bezerra, pós-graduada pelo Programa de Pós-Graduação em Letras da Universidade Federal de Pernambuco (UFPE), que realizou a coleta de 80 redações em 2018 e me concedeu o acesso a elas por meio da aprovação do projeto de pesquisa por mim submetido ao Comitê de Ética em Pesquisa da UFPE (CEP-UFPE) (nº do processo: CAAE nº 23247719.9.0000.5208). De acordo com a pesquisadora, os critérios de seleção das redações envolveram: a) redações com maiores e menores notas obtidas na avaliação do texto pela UNILAB e b) redações produzidas por homens e mulheres (BEZERRA, 2020). Esses critérios permitiram conferir tanto variedade como paridade aos dados analisados a partir dos textos.

3.3 PARTICIPANTES DA PESQUISA

Os participantes da pesquisa que produziram as redações são timorenses natos, todos falantes do tétum e do português e já concluintes da educação básica (o que equivale ao atual Ensino Médio no Brasil). Guimarães (2017, p. 506) assegura que em Timor-Leste o português é aprendido por meio do tétum em todo o processo de escolarização. Ele diz:

> Em consonância com a Constituição do país, a Lei de Bases da Educação de Timor-Leste (n.º 14/2008) em seu Artigo 8.º aponta: "As línguas de ensino do sistema de ensino timorense são tétum e o português." Ou seja, o ensino de LP deve ocorrer junto ao ensino do tétum, sendo esta língua utilizada para auxiliar o ensino daquela.

Desse modo, é razoável supor que todos os sujeitos da pesquisa, uma vez escolarizados, também são falantes do tétum[16]. Além disso, a fim de garantir uma simetria com base no sexo dos informantes, foram selecionadas 80 redações, sendo 40 produzidas por homens e 40, por mulheres.

3.4 CONSTITUIÇÃO DO *CORPUS*

O *corpus* é constituído de um total de 582 dados que envolvem sentenças declarativas finitas em que estão presentes sujeitos nulos ou plenos. Para tanto, foi realizada uma atenta leitura das 80 redações na procura e identificação desses sujeitos. A seguir, observem-se alguns exemplos de frases extraídas do *corpus*:[17]

(69) *pro* acho que *pro* preciso contribuir as idéias na universidade (T1, M)
(70) *Timor-Leste* faz parte da comunidade países lingua português (T6, M)
(71) *A lingua portuguêsa* é a lingua oficial (T11, M)

[16] Infelizmente não foi possível entrevistar ou cadastrar cada informantes para saber qual a sua LM ou outras línguas que podem falar. Desse modo, evitarei fazer conclusões ao longo desta pesquisa sobre a influência específica do tétum. No entanto, é válido considerar que o tétum é, ao menos, conhecido de todos os informantes, por terem obtido instrução formal em Timor-Leste, instrução ministrada unicamente em tétum.

[17] A informação entre parênteses se refere ao número do texto (T) e ao sexo do escrevente: masculino (M) e feminino (F). Decidi manter, neste trabalho, a grafia dos exemplos tal como ocorre nos textos. Não foi feita, portanto, nenhuma correção, atualização ou verificação ortográfica.

(72) *pro* quero muito de desenvolver minha capacidade (T63, F)
(73) *A nação de Timor-Leste* têm a necessidade de insistir na Língua Portuguesa (T65, F)

Excluí da constituição do *corpus*, à semelhança do que fez Duarte (1995), construções impessoais (a saber: construções com os verbos *haver* e *ter* no sentido de *existir* e construções verbais relacionadas a fenômenos da natureza), bem como construções encaixadas no infinitivo impessoal.

O quadro a seguir esquematiza a constituição do *corpus* desta pesquisa.

Quadro 10 - Constituição do *corpus*

Informantes da pesquisa	Estudantes timorenses	
Contexto da elaboração do *corpus*	Escrita de redações para ingresso no ensino superior	
Local de produção das redações	Timor-Leste	
Local de coleta das redações	UNILAB (Redenção-CE)	
Total de redações coletadas	80	
	40 escritas por homens	40 escritas por mulheres
Total de dados extraídos das redações	582	

Fonte: autor deste trabalho

3.5 SELEÇÃO DAS VARIÁVEIS

Para a realização da análise linguístico-quantitativa, selecionei variáveis (linguísticas e extralinguísticas). De maneira geral, o leque de variáveis elegidas neste estudo foi baseado nas variáveis já analisadas por Duarte (1995) em seu estudo pioneiro sobre o PSN no PB. Além disso, entendi como proveitoso selecionar para a análise a posição do sujeito pleno, tendo em vista o fato de ser uma propriedade das

línguas *pro-drop* a ocorrência de sujeitos pospostos, como ocorre no PE (COSTA, 1998), variedade que é ensinada em Timor-Leste.

3.5.1 Variáveis

3.5.1.1 Variável dependente

1. *Sujeito nulo*

(74) a. *pro* queria ser uma boa puncionario publico (T23, M)
b. *pro* gosto muito de aprender língua portuguesa (T63, F)
c. *pro* pretendo cursar na area de enfermagem (T66, F)

2. *Sujeito pleno*

(75) a. *a lingua portuguesa* é a chave da nação (T21, M)
b. *Timorenses* usa a lingua portuguesa (T63, F)
c. *Timor-Leste* é uma nação que muito pequena (T67, F)

3.5.1.2 Variáveis linguísticas

- Realização do sujeito pleno

1. *Pronome*

(76) a. *Eu* vou de curso de UNILAB (T61, F)
b. *Nós* usamos lingua português para fazir a politica diplomasica (T31, M)

2. *Sintagma Nominal*

(77) a. *a lingua português* é muito importante para os timorenses (T33, M)
b. *o Timor-Leste* têm vários recursos naturais (T76, F)

Optei, nesta pesquisa, pela inclusão de sujeitos realizados por sintagmas nominais, o que não está previsto em Duarte (1995). Essa inclusão ocorre por dois motivos: (1) a possibilidade de mais contextos declarativos para a análise, por meio dos quais será possível analisar aspectos relacionados à natureza da morfologia de flexão verbal (AGR) e à posição e duplicação do sujeito pleno, e (2) a expectativa prévia de amplo uso de sujeitos dessa natureza em virtude do contexto de produção linguística envolvido: escrita formal monitorada, em que um dado tópico precisa ser desenvolvido pelo escrevente, sendo retomado ao longo do discurso. De toda forma, os resultados obtidos serão apresentados de maneira separada: a) para a análise da concordância verbal (ou seja, da morfologia de flexão verbal), serão considerados todos os sujeitos plenos (sintagmas nominais e pronomes) e, para análise da distribuição de sujeitos nulos e plenos, serão apenas incluídos nestes últimos os sujeitos pronominais, assim como fizera Duarte (1995).

- Posição do sujeito pleno

1. *Posição pré-verbal*

(78) a. *Os Portugueses ensinam* muitas coisas (T58, F)
b. *Ensino superior público* utiliza a língua Indonesio (T52, F)

2. *Posição pós-verbal*[18]

A variável posição do sujeito pleno não foi abordada por Duarte (1995) em sua tese. No entanto, outras pesquisas posteriores demonstraram que a ordem dos elementos expressa movimentos. As línguas *não-pro-drop* tendem a manter cada vez mais rígida a ordem SVO (SILVA, 2004). Isso acontece também em virtude da simplificação dos paradigmas de concordância dessa língua, como mencionarei adiante. Silva (2004) também verificou que no PE é possível, além dessa ordem, o

[18] Não foram encontradas ocorrências no *corpus*.

sujeito posposto ao verbo com todos os tipos de verbos, o que é característico de uma língua de sujeito nulo consistente. Portanto, ao incluir essa variável, tive o intuito de observar como o PTL se comporta no que se refere à inversão da ordem S-V. À revelia do confronto quantitativo ensejado, a manifestação categórica dessa variável nos dados analisados foi pela não ocorrência de sujeitos na posição pós-verbal.

- Tipo de verbo

1. *Intransitivo*[19]

2. *Transitivo*

(79) a. os cidadão dos país *utilizam* a lingua Tétum (T4, M)
b. eu *queria* expressar o que é que eu aprendo (T44, F)

3. *Inacusativo*

(80) a. População de Timor-Leste *vive* numa economia agrícola (T2, M)
b. Quando *voltamos* para o nosso pais (T62, F)

4. *Copulativo*

(81) a. A lingua portuguesa *é* uma lingua que nós uzamos (T52, F)
b. Lingua português *é* importante para os timorenses (T34, M)

Os fatores que compõem a variável em análise foram definidos tomando por base o que consta em Duarte (2003, p. 296-303). Sendo assim, define-se o verbo intransitivo como um predicador de um lugar cujo argumento externo tem relação gramatical de sujeito; transitivo, o predicador de dois ou mais lugares cujo argumento externo tem relação gramatical de sujeito; inacusativo, o predicador que seleciona argumento interno ao qual não é atribuído Caso acusativo e que exerce a relação gramatical de sujeito; copulativo, aquele que seleciona uma oração pequena.

[19] Não foram encontradas ocorrências no *corpus*.

Segundo Duarte (2003, p. 303), “superficialmente, o sujeito da oração pequena ocorre com a relação gramatical de sujeito da frase copulativa”.

- Tipo de oração

1. *Oração encaixada*

(82) a. eu vejo *que os timorenses usam* so na escola (T24, M)
b. o país declarou *que português é lingua oficial* (T33, M)

2. *Outras orações*

(83) a. *Timor Leste decidiu o indonesia e inglês* (T38, M)
b. *lingua portuguêsa é um língua mais falado* (T60, F)

Os contextos de oração encaixada tendem a favorecer os nulos, especialmente se o sujeito dessa oração mantém correferência com o sujeito da oração principal. Duarte (1995, p. 43) afirma que as encaixadas são um fator fundamental, pois, em línguas como o italiano, que são *pro-drop*, “estruturas subordinadas com sujeitos correferentes têm obrigatoriamente o segundo ou ambos os sujeitos nulos.”

Similarmente, Holmberg, Nayadu e Sheehan (2009) afirmam que o sujeito nulo em orações encaixadas complementos também são características de línguas de sujeito nulo parcial, como o PB. A mesma observação foi feita por Kato e Duarte (2014), constatando a alta produtividade de sujeitos nulos em orações encaixadas no PB. Desse modo, a presença dessa característica do PB mostra que ele não é uma língua em que o sujeito nulo é agonizante, assim como afirmara Duarte (1995). Por outro lado, também não é uma língua de sujeito nulo consistente tal qual o PE.

Relacionado a isso, ainda que seja apontada como *não-pro-drop*, a língua tétum permite o sujeito nulo em contextos restritos. Por exemplo, há o nulo expletivo. Além disso, Sofia Deus (2011) mencionou que o tétum admite sujeitos referenciais nulos em orações encaixadas em que não há o uso de quantificadores, como ninguém ou algum. Em tétum, ocorre nulo em “quem pensa que ele é esperto, *pro* levante a mão” (DEUS, 2011, p. 237). Como os nulos em tétum acontecem nas orações

encaixadas, mesmo que de maneira mais restrita em relação ao PE e ao PB, acredito ser válido observar como o PTL se comporta nesse contexto oracional.

- Duplicação do Sujeito Pleno

1. *Presença de duplicação*[20]

2. *Ausência de duplicação*

(84) a. *Norma Culta* é o conjunto de práticas linguisticas (T26, M)
b. *o português* entra na nação de Timor-Leste (T19, M)

Duarte (1995) selecionou a duplicação do sujeito como variável, contexto em que um sintagma nominal é retomado por um pronome com os mesmos traços número-pessoais, conforme exemplos a seguir:

(85) a. *A Clarinha ela* cozinha que é uma maravilha.
b. Eu acho que *o povo brasileiro ele* tem uma grave doença.
c. Não vou falar de bermuda, porque *os alunos* hoje em dia no verão *eles* vêm assistir aulas com bermuda de qualquer tamanho.
(DUARTE, 1995, p. 108)

Segundo a autora, construções com duplicação do sujeito são uma das estratégias de preenchimento da posição sujeito no PB, em especial, na língua falada. Essas construções são “conhecidas como deslocamento à esquerda ou à direita [e] são construções atípicas em línguas *pro-drop* e... seu exame é fundamental para a confirmação da hipótese de distanciamento do grupo das línguas *pro-drop*.” (DUARTE, 1995, p. 44).

De antemão, é possível afirmar, desde já, que os dados do PTL revelam um comportamento interessante, pois sujeitos duplicados não foram encontrados, o que

[20] Não foram encontradas ocorrências no *corpus*.

talvez se deva ao monitoramento da modalidade escrita da língua ou uma aproximação maior, nesse sentido, ao PE, língua tipicamente *pro-drop*.[21]

- Pessoa e número do sujeito

1. *1ª pessoa do singular*

(86) a. *Eu* gosto de falar e estudar (T66, F)
b. *pro* acho que português é importante (T21, M)

2. *2ª pessoa do singular*[22]

3. *3ª pessoa do singular*

(87) a. *Timor-Leste* é uma nação que muito pequena (T67, F)
b. *a lingua português* é a oficial de Timor-Leste (T31, M)

4. *1ª pessoa do plural*

(88) a. *nós* estudamos a língua portuguesa (T62, F)
b. *Nós* vimos para a historia de Timor-Leste (T31, M)

5. *2ª pessoa do plural*[23]

6. *3ª pessoa do plural*

(89) a. *Os estudantes de Timor-Leste* tem obrigatório para aprender (T38, M)
b. *Os agradecimentos em portugues* são parte do dia a dia (T53, F)

[21] Como se verá no capítulo seguinte, os dados aqui em análise não apresentaram duplicação do sujeito. No entanto, isso não significa afirmar categoricamente que no PTL não ocorre essa duplicação. Portanto, estudos futuros são necessários sobre essa variedade do português seja na língua escrita seja na língua falada para verificar se contextos de duplicação do sujeito são produzidos.
[22] Não foram encontradas ocorrências no *corpus*
[23] Não foram encontradas ocorrências no *corpus.*

Duarte (1995) observou que os contextos de terceira pessoa, tanto no singular como no plural, foram mais relevantes para ocorrência do sujeito nulo no PB. Desejo observar, por sua vez, se a referência semântica do sujeito tem significância no PTL para a produção de sujeitos nulos.

- Morfologia de flexão verbal

Para essa variável, foram considerados os casos em que ocorre a concordância padrão e a concordância não-padrão entre os sujeitos realizados foneticamente e a flexão verbal.

1. *Presença de concordância padrão*

(90) a. *eu quero* desenvolver a nosso nação (T11, M)
b. *Nós podemos* comunicar com as pessoas de estrangeiro (T26, M)

2. *Ausência de concordância padrão*

(91) a. *todas as populações falamos* em lingua português (T43, F)
b. Os timorense precisa aprender (T19, M)
c. *eu aproveita* têmpo para studar (T25, M)

Os estudos pioneiros sobre o PSN defendiam que a morfologia de flexão (AGR) rica licencia e identifica os sujeitos nulos (CHOMSKY, 1981, RIZZI, 1986). O PB, conforme já dito, tem passado por um enfraquecimento de AGR e ainda produz sujeitos nulos (DUARTE, 1995, GALVES, 2001). Em sua tese, Duarte (1995) afirmou que a escolha dessa variável foi “fundamental para mostrar que, cada vez menos, é a flexão o elemento que licencia o sujeito nulo e que isso ocorre, pela perda da riqueza funcional” (p. 41).

Conforme apresentei no capítulo 1 desta dissertação, o Tétum não tem morfemas número-pessoais que identifiquem as pessoas do paradigma e, por isso, o AGR é pobre nessa língua. Diante disso, levanto a seguinte questão: os timorenses podem preferir sujeitos plenos ao aprenderem o português e produzirem menos concordância? Como o acesso à GU para a aprendizagem de português se dá pela(s) língua(s) já adquirida(s), como defendo, seria provável observar a gramática

internalizada do timorense dificultando a marcação paramétrica no PTL, tornando essa variedade afastada da variedade europeia do português. Sendo assim, por ser uma LE, é provável que os sujeitos se comportem diferentemente da norma europeia do português, no sentido de que haja menos nulos no PTL.

3.5.1.3 Variável extralinguística

- Sexo do escrevente

1. *Feminino*
2. *Masculino*

A escolha dessa variável justifica-se, tal como apontado por Duarte (1995), pelo fato de as mulheres produzirem menos sujeitos nulos no PB. Labov (1972) propôs que no processo de mudança linguística, as mulheres evidenciam as novas formas antes que os homens. Duarte, que até então defendia uma mudança no parâmetro em português, argumentou que as mulheres tendem a ser inovadoras nesse contexto. Ela afirmou:

> Ora, sabe-se que as mulheres tendem a liderar um processo de mudança em direção a uma forma prestigiada (Paiva, 1992), mantendo-se mais conservadoras em relação a formas não prestigiadas. Como no caso da presença ou ausência do pronome sujeito não se tem variantes sujeitas a estigma social, pode-se igualmente supor que as mulheres "tomem a ponta do processo de mudança" (Paiva, 1992, p. 71). (DUARTE, 1995, p. 50)

No caso de Duarte (1995), foi exatamente isso o que ocorreu: as mulheres produziram menos sujeitos nulos que os homens. Estes demonstraram, em valores percentuais, 34% de sujeito nulo, ao passo que aquelas produziram 25% de sujeitos nulos. Levando em conta esse caráter inovador das mulheres, no PTL, se houver um processo de aprendizagem mais próximo da variedade europeia, esperar-se-á que as mulheres produzam mais sujeitos nulos. No entanto, se houver um distanciamento dessa variedade, as mulheres tenderão a produzir menos sujeitos nulos.

O quadro a seguir apresenta os fatores controlados por meio dos quais os dados do *corpus* foram codificados e submetidos à análise:

Quadro 11 - Fatores controlados das variáveis

Variável	Fatores
Tipo de sujeito	Nulo
	Pleno
Realização do sujeito pleno	Pronome
	Sintagma nominal
Tipo de verbo	Intransitivo
	Transitivo
	Inacusativo
	Copulativo
Tipo de oração	Encaixada
	Outras orações
Pessoa e número do sujeito	1ª pessoa do singular
	3ª pessoa do singular
	1ª pessoa do plural
	3ª pessoa do plural
Morfologia de flexão verbal	Presença de concordância padrão
	Ausência de concordância padrão
Sexo do escrevente	Feminino
	Masculino

Fonte: Autor deste trabalho.

3.6 TIPO DE ANÁLISE

Após a codificação dos dados, todos foram submetidos a tratamento quantitativo através do programa GoldVarb X (SANKOFF; TAGLIAMONTE; SMITH, 2005). O uso desse programa computacional possibilitou-me a obtenção de resultados totais e percentuais. De modo geral, o GoldVarb X é utilizado em pesquisas

variacionistas que procuram dar conta dos fatores que favorecem ou desfavorecem o uso de uma determinada forma linguística. Nesse caso, o programa permite ter acesso também aos pesos relativos que revelam o nível de significância desses fatores.

Esses pesos relativos são a indicação do efeito que cada fator selecionado exerce sobre a variável em análise. O fator tem um maior nível de significância se seu valor for superior a 0.50 e menor nível de significância se for inferior a 0.50 (SCHERRE; NARO, 2012). Assim, o GoldVarb X gera “como produto final, resultados numéricos associados aos diversos fatores dos grupos de fatores (variáveis), que medem o efeito relativo de cada fator” no fenômeno sob análise (SCHERRE; NARO, 2012, p. 161). E ainda, “os resultados numéricos obtidos têm só valor estatístico. O seu valor linguístico é atribuído e interpretado pelo linguista” (SHERRE; NARO, 2012, p. 162).

Apresentadas essas características atreladas ao uso do GoldVarb X pela pesquisa sociolinguística variacionista, informo que nesta pesquisa me limitei ao uso do programa para obtenção dos valores quantitativos totais (número de ocorrências) e dos percentuais, de modo que não pretendo abranger os pesos relativos. Nesse sentido, a análise apresentada, no próximo capítulo, é de natureza linguístico-quantitativa.

4 COMPREENDENDO A CONCORDÂNCIA E O SUJEITO NULO NOS DADOS ESCRITOS POR ESTUDANTES TIMORENSES

Neste capítulo, apresentarei e analisarei o comportamento da concordância verbal e as propriedades do sujeito nulo nos dados de escrita do PTL a partir dos resultados quantitativos obtidos no GoldVarb X. Também pretendo explorar a realização de formas verbais infinitas com sujeitos plenos e o que esse resultado possibilita compreender da aprendizagem do português pelos timorenses.

Durante toda a análise, tendo em vista o fato de o PTL estar inserido em um contexto multilíngue e de o contexto de produção escrita dos dados favorecer um maior monitoramento dos timorenses quanto à maneira como estão escrevendo, objetivo discutir se nessa escrita ainda é possível verificar evidências de distanciamento com relação ao que é encontrado na variedade europeia do português no âmbito do PSN. De acordo com Souza et al. (2010), a escrita é menos dada a inovações do que a fala, sobretudo, quando se trata de contextos mais monitorados, nos quais o texto pode passar por várias etapas de reescrita até a sua versão final. Ciente disso, avento a hipótese de que, mesmo havendo a possibilidade de reescrita em um contexto de produção monitorado, é possível encontrar construções linguísticas que fogem ao que se observa no PE em relação ao PSN por razões relacionadas à maneira como o PTL, que está inserido em um contexto multilíngue e sofre interferência deste, é aprendido como uma LNM pelos falantes.

Este capítulo, portanto, está dividido em quatro seções. Primeiramente, farei a apresentação dos resultados gerais incluindo sujeitos nulos e sujeitos plenos realizados sob a forma de sintagmas nominais e pronomes. Em seguida, delimitarei a análise, enfocando apenas os sujeitos nulos e os sujeitos pronominais, tal como foi desenvolvido por Duarte (1995). Após isso, será feita uma discussão dos resultados e, por fim, apresentarei uma breve incursão nas construções com verbos no infinitivo encontradas nas redações analisadas.

Inicialmente, elenco, a seguir, os resultados gerais obtidos, em valores percentuais, para o processamento das variáveis controladas relacionadas para a

análise, com o auxílio de tabelas e gráficos que permitam uma visão geral da distribuição dos sujeitos no *corpus*.

Pretendo, ainda, dialogar com alguns dos trabalhos já mencionados nos capítulos anteriores, em especial, Duarte (1995), Kato e Duarte (2014) e Malta (2019) para tecer considerações comparativas com os resultados deste trabalho em relação ao sujeito nulo em outras variedades do português. Trarei também os trabalhos de Santos (2009) e Martins (2016) sobre o PTL para observar os seus resultados e as conclusões a que chegaram no exame de textos escritos, tal qual aqui tenho me empenhado em fazer também.

4.1 RESULTADOS GERAIS

Para os resultados gerais, foram selecionados 582 dados. Ou seja, todas as sentenças finitas em que poderiam ocorrer sujeitos plenos (incluindo, neste primeiro momento, os sintagmas nominais e os sujeitos pronominais) e nulos.

4.1.1 Resultados por variável

4.1.1.1 Tipo de verbo

Dentre os tipos de verbos analisados no total de dados selecionados, percebi uma quantidade maior de copulativos e transitivos. Juntos, esses dois tipos verbais representam mais de 95% do total de ocorrências, conforme pode ser observado no quadro 12.

Quadro 12 - Totais por tipo de verbo nos dados gerais

Tipo de verbo	Ocorrência no *corpus*	%
Copulativo	226/582	39,8%
Transitivo	339/582	58,4%
Inacusativo	16/582	2,8%

Fonte: Autor deste trabalho

a) Verbos copulativos

(92) a língua portuguesa até agora *é* muito fácil (T43, F)
(93) o país ainda *é* novo (T46, F)
(94) A Lingua português é muito importante (T16, M)
(95) porque *pro é* uma língua oficial da nossa nação (T5, M)
(96) *pro são* pessoas atender para família (T48, F)

b) Verbos transitivos

(97) Man Alkalin *declarou* que a implementação... (T12, M)
(98) as minhas família sempre *falam* a lingua portuguêse (T8, M)
(99) meus avós *utilizam* os Língua português (T20, M)
(100) *pro queria* implementar as ciências e experiências que aprender (T21, M)
(101) *pro acho* que é importante colabora (T27, M)

c) Verbos inacusativos

(102) a lingua que *existem* em Timor-Leste (T6, M)
(103) População de Timor-Leste vive numa economia agrícola (T2, M)
(104) quando *pro* voltamos para o nosso pais (T62, F)
(105) *pro* [o curso] começa no ano 2009 (T16, M)
(106) *pro* [os timorenses] vão ao Brazil para saber (T37, M)

4.1.1.2 Tipo de oração

Poucas orações foram encaixadas, apenas 18 de 582. De fato, durante todo o levantamento de dados feito nas redações que compuseram a pesquisa, algo que chamou a atenção foram as orações, em sua maioria, independentes, com apenas um verbo, e períodos curtos.

Quadro 13 - Totais por tipo de oração nos dados gerais

Tipo de oração	Ocorrência no *corpus*	%
Encaixada	18/582	3,1%
Outras orações	564/582	96,9%

Fonte: Autor deste trabalho

a) Orações encaixadas

(107) eu acho *que a matéria da lingua portuguêsa é muito importante* (T69, F)
(108) *pro* acho *que preciso contribuir as idéias na universidade* (T1, M)
(109) eu vejo *que os timorenses usam so na escola* (T24, M)
(110) pro quero *que todo mundo conhece Timor-Leste (T68, F)*
(111) o país declarou *que português é lingua oficial* (T33, M)

b) Outras orações

(112) *pro queria ajudar outros nação* (T76, F)
(113) *pro consideramos como lingua oficial* (T78, F)
(114) *a língua está usar na educação primária* (T38, M)
(115) *Timor Leste é um país soberano* (T66, F)
(116) *País Timor-Leste são país independentes* (T79, F)

4.1.1.3 Pessoa e número do sujeito

No *corpus* analisado, não houve ocorrências de referência semântica do sujeito para a segunda pessoa, seja do singular, seja do plural. O próprio contexto de produção desses textos, que era formal e não envolvia interlocução entre dois participantes no momento concreto da enunciação, não favoreceu essa produção.

Um fato curioso é que foi encontrado no *corpus* um único uso do pronome *você* com leitura genérica, não definida:

(117) pois se *você* ter uma boa base educação boa, sólida, uma situação boa, confortáveis, *você* consegue reproduzir melhor e aprende (T44, F)

Em sua tese, Duarte (1995) dedicou uma seção para os sujeitos de referência não definida, que foram muito produtivos em sua pesquisa com o PB, ao contrário do que foi observado no PTL aqui investigado. Uma possível motivação para a diferença é que, como a pesquisa de Duarte envolveu dados de fala e este estudo volta-se à análise de dados de escrita, os sujeitos de referência não definida podem ter maior produtividade mais naquela modalidade do que nesta.

Conforme o quadro 14, a seguir, a maioria dos sujeitos no *corpus* foi de terceira pessoa do singular, 56,1%.

Quadro 14 - Totais por tipo de referência semântica do sujeito nos dados gerais

Pessoa e número	**Ocorrência no *corpus***	**%**
1ª singular	152/582	26,2%
3ª singular	326/582	56,1%
1ª plural	25/582	4,3%
3ª plural	76/582	13,1%

Fonte: autor deste trabalho

a) 1ª pessoa do singular

(118) *pro acho* que é importante colabora (T27, M)
(119) *pro quero* ajudar as crianças (T46, F)
(120) *eu pretendo* cursar uma universidade (T48, F)
(121) *eu sou* uma nova geração (T77, F)

b) 3ª pessoa do singular

(122) *pro* [A minha experiência] *pode* falar muito a lingua portuguêsa (T11, M)
(123) *pro* [A educação] *Começa* a ensinar as matérias (T16, M)
(124) *Timor-Leste tem* dois Lingua portugues e tetum (T25, M)
(125) *Brazil é* alguma nação que também usar esse língua (T27, M)

c) 1ª pessoa do plural

(126) *pro fazemos* a relação com multilateral (T26, M)
(127) *pro precisamos* tambem a firmação da língua portugueses (T35, M)
(128) *Nós podemos* fazer economia (T26, M)
(129) *nós começamos* com a Lingua portuguesa (T19, M)

d) 3ª pessoa do plural

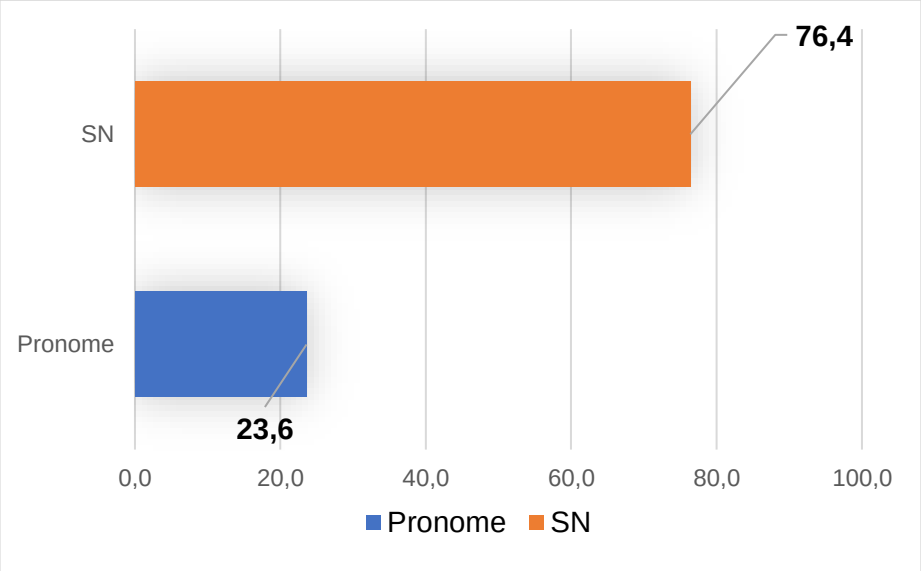
(130) *pro* [Os estudantes] *voltaram* pra nosso pais (T16, M)
(131) *pro* [Muitos professores] *utilizam* a lingua tétum (T43, F)
(132) *os professores ensinam* as materie com lingua tetum (T8, M)
(133) *o indonésio e o tetum são* a lingua de meio ambiente (T12, M)

4.1.1.4 Realização do sujeito pleno

Os resultados, a seguir, foram obtidos com base nos 471 sujeitos plenos do total de dados. Dos 471 sujeitos plenos do *corpus*, 111 deles foram expressos por pronomes (23,6% do total) e 360 por sintagmas nominais (76,4%), conforme

apresentado no Gráfico 1. Essa alta produtividade de sujeitos plenos por sintagmas nominais ocorreu conforme o esperado, em vista do contexto de produção e do monitoramento da escrita que ele exige. Os tópicos apresentados são retomados ou repetidos durante o desenvolvimento do texto, como recursos textuais de coesão (ANTUNES, 2005). No caso dos alunos timorenses, foi recorrente a repetição de sintagmas como “língua portuguesa”, “português” e “Timor-Leste”.

Gráfico 1 - Realização do sujeito pleno (%)



Fonte: autor deste trabalho

a) Sujeito pleno expresso por pronome

(134) *eu* quero criar uma curso (T2, M)
(135) *eu* quero aprender mais lingua portuguesa (T3, M)
(136) *Eu* acho importante para colaborar (T7, M)
(137) *eles* sempre falaram a lingua portuguesa (T8, M)
(138) *eu* vou continuar para terminar a minha escola (T10, M)
(139) as capacidades que *eu* tenho pra colaborar (T37, M)
(140) *eles* são graduado no tempo da Indonésia (T39, M)
(141) *eu* deverei estudar no estrangeiro (T46, F)
(142) *nós* usa lingua portuguesa (T49, F)
(143) *Nós* entendemos mais fácil (T63, F)

b) Sujeito pleno expresso por sintagma nominal

(144) *os cidadão dos país* utilizam a lingua Tétum (T4, M)

(145) *Timor-Leste* tem muitos doutores e enfermeiros (T5, M)

(146) *Tetum* é lingua nacional (T6, M)

(147) *os professores* ensinam as materie com lingua tetum (T8, M)

(148) *Timor-Leste* é uma nação democrática (T9, M)

(149) *o Indonesio e o tetum* são mais frequentes nos ambientes de trabalho (T53, F)

(150) *os timorenses* falam com a língua Tétun (T63, F)

(151) *o nosso pais* é novo pais (T64, F)

(152) *as comunidades* queria falar bem (T65, F)

Tendo em vista que o preenchimento da posição sujeito em português pode dar-se tanto por meio de pronomes não expletivos como por sintagmas nominais movidos à posição Spec,IP, é necessário compreender o contexto de produção do material linguístico para a adequada apresentação dos resultados quantitativos obtidos.

Em um texto escrito produzido em um contexto de maior monitorado, como é a redação de um exame de seleção para ingresso em uma universidade, o escrevente deve demonstrar conhecimento, durante a sua formação, das estratégias de coesão textual. Antunes (2005) aponta que, dentre os vários procedimentos que conferem ao texto coesão, está a *reiteração* entendida como a retomada de referentes textuais que agem como tópicos discursivos, colaborando assim tanto para a manutenção como para a progressão temática.

Os pronomes podem ser uma excelente estratégia de retomada, como *meu país > ele,* mas têm limitações no contexto de uso. Espera-se de um escrevente, ao escrever um texto dissertativo que se materializa em uma redação, que evidencie várias estratégias/relações coesivas (reiteração, substituição, conexão e associação (cf. ANTUNES, 2005)) e demonstre variedade vocabular.

Na prática, é esperado que os sintagmas nominais sejam usados para auxiliar no processo de coesão e construção textual. Pelo alto grau de monitoramento a que se submetem durante a escrita desse texto, os escreventes podem “corrigir” a escrita de um pronome ou sintagma nominal repetido por outro sintagma nominal que seja textualmente adequado à sua produção.

Em uma pesquisa anterior (OLIVEIRA JÚNIOR, 2018), analisando redações de alunos produzidas durante um exame vestibular no PB, deparei-me com a mesma situação, em que a quantidade de pronomes preenchendo a posição sujeito (17,3%) foi consideravelmente menor que a quantidade de sintagmas nominais (82,7%). No que se refere ao PTL, há um contexto de produção similar e, por isso, as ocorrências expressivas de SN como sujeito pleno

4.1.1.5 Posição do sujeito pleno

Para essa variável, todos os sujeitos plenos do *corpus* foram pré-verbais. Isso evidencia uma estrutura rígida na ordem Sujeito-Verbo. Essa ordem, que costuma ser passível de inversão para Verbo-Sujeito em línguas *pro-drop* como o PE (NICOLAU, 1995; SILVA, 2004), se comporta de maneira diferente no PTL. Por outro lado, o enrijecimento da ordem S-V é mais característico de línguas *não-pro-drop*, como o tétum.

A ordem de palavras tem sido alvo de muitos estudos gerativistas para o arcabouço descritivo das línguas. Sofia Deus (2010) e Albuquerque (2011) já mencionavam que o tétum tem estrutura SVO. Isso parece ser também uma característica de outras línguas do tronco malaio-polinésio, visto que Randriamasimanana (2019) analisou cerca de três dezenas de línguas malaio-polinésias e apresentou que tanto o malaio como o indonésio, línguas de falantes que circunvizinham Timor-Leste, têm ordem Sujeito–Verbo–Objeto. Os exemplos a seguir são ilustrativos nesse respeito:

(153) Paul memotong kayu. [Bahasa Indonesia]
S V O
'Paul chopped the wood.'

(RANDRIAMASIMANANA, 2019, p. 266)

(154) hau sei halo kafee mai ami. [Tetum-Dili]
S V O
‘Eu farei café para nós’

(ALBUQUERQUE, 2011, p. 109)

Embora não seja conclusivo, é possível aventarmos a hipótese de que, uma vez que o PE é ensinado em Timor Leste, seria esperado que houvesse inversão da ordem SV em algum momento no texto, o que não se confirma, distanciando o PTL do PE.

4.1.1.6 Duplicação do sujeito pleno

Para essa variável em análise, também não houve ocorrências no *corpus*. Segundo estudiosos, a duplicação do sujeito resulta de movimento do sujeito para uma posição de Tópico, a chamada Construção de Deslocamento à Esquerda do sujeito, que é retomado por um pronome (pronome cópia, pronome resumptivo) no domínio intrafrásico na posição Spec, IP (ex.: *A Raquel, ela* canta muito.). Em sua pesquisa, Duarte (1995) observou ocorrências de duplo sujeito na língua falada do PB e concluiu que a duplicação é uma estratégia muito produtiva de preencher a posição sujeito nessa língua.

No caso do PTL, mesmo com um alto índice de sujeitos plenos realizados por sintagmas nominais, nenhum deles foi retomado por um pronome cópia. Esse resultado pode ter grande relação com o contexto de monitoramento da escrita. Por exemplo, não encontrei nenhum caso de duplicação do sujeito na escrita monitorada do PB quando analisei também redações escritas por estudantes de nível médio para ingresso no ensino superior (OLIVEIRA JUNIOR, 2018). Sendo assim, não é possível afirmar que no PTL não exista a duplicação do sujeito, tomando por base apenas os dados de escrita. Portanto, faz-se necessária a realização de pesquisa futura na língua falada dessa variedade para que se possa verificar se a duplicação do sujeito é produzida nessa modalidade de uso da língua e se serve como estratégia de preenchimento da posição pré-verbal do sujeito, tal como verificado na fala do PB por Duarte (1995).

4.1.1.7 Morfologia de flexão verbal

O PTL apresentou, no *corpus* analisado, maior produção de concordância padrão (78,3%). Ainda que esteja inserido em um contexto de multilinguismo em que as línguas faladas localmente, a exemplo do tétum, não têm morfologia de flexão verbal (DEUS, 2011), observo a inegável influência do processo de ensino do PE em Timor Leste motivando a grande frequência de uso da concordância padrão.

Retomando o resultado de Brandão e Vieira (2012, p. 1040), de que, na língua falada do PE urbano, a concordância verbal é semicategórica (98,9%), ao passo que, na língua falada do PB, é variável (78,1%), o comportamento da concordância verbal na língua escrita do PTL, com o total de 78,3%, é muito próximo ao da língua falada do PB, conforme apresentado no quadro 15 a seguir:

Quadro 15 - Concordância verbal do sujeito pleno

Concordância	Ocorrência no *corpus*	%
Padrão	369/471	78,3%
Não padrão	102/471	21,7%

Fonte: autor deste trabalho

a) Concordância padrão

Nos casos de concordância padrão a seguir apresentados, serão vistos verbos junto a sujeitos de representação semântica tanto para as pessoas do singular como do plural. Na perspectiva da realização desta pesquisa, em que o conjunto de dados contribui para o entendimento mais acertado do comportamento da variedade timorense do português, os sujeitos no plural foram poucos. Além disso, como houve casos de concordância não-padrão com o sujeito de referência semântica no singular e verbo no plural, defini que a concordância ocorre quando as marcas semântico-discursivas do sujeito e do verbo são correspondentes:

(155) *português é* lingua internacional (T19, M)
(156) *nós começamos* com a Lingua portuguesa (T19, M)
(157) *o português entra* na nação de Timor-Leste (T19, M)

(158) *a lingua portuguesa é* a chave da nação (T21, M)
(159) *a Lingua portuguesa é* Lingua oficial de Timor-Leste (T24, M)
(160) *eu quero* estabelecer na minha conhecimento (T26, M)
(161) *Nós usamos* lingua português para fazir a politica diplomasica (T31, M)
(162) *Eu agradeco* pelo UNILAB (T35, M)
(163) *A língua portuguesa é* muito importante para aprender (T47, F)
(164) *O governo decidiu* que o idioma veicular da educação seja a lingua portuguesa (T49, F)

b) Concordância não-padrão

(165) *eu aproveita* têmpo para studar (T25, M)
(166) *os timorenses preciso* de conhecer (T30, M)
(167) *Timor-Leste adaptam* a lingua português (T31, M)
(168) *nós utiliza* língua no país (T41, F)
(169) *nós usa* lingua portuguesa (T49, F)
(170) *o recursos humanos é* um fator muito importante (T55, F)
(171) *Alkatiri têm* outra razão (T57, F)
(172) *Timor-Leste usam* de lingua portuguêsa (T61, F)
(173) *a nação Timor-Leste são* membro da CPLP (T67, F)
(174) *o tétum e português é* como a língua oficial (T71, F)

Vale referir que, a respeito do fator escolaridade interferindo no comportamento da concordância, é comum que pessoas com menor nível de escolarização produzam menos concordância. Estruturas como “Eu canta” foram verificadas por Mota, Miguel e Mendes (2012) em falantes pouco escolarizados de algumas variedades africanas no português (ex.: o PST e o PA). Não obstante, o interessante aqui é que os timorenses já têm um grau de escolarização alto, estão produzindo um texto na modalidade escrita que exige um maior grau de monitoramento e, mesmo assim, há casos de concordância não-padrão similares aos que foram verificados por Mota, Miguel e Mendes (2012), o que aponta para um enfraquecimento de sua morfologia de flexão verbal. Isso me leva a argumentar que a dificuldade de aprender a morfologia de flexão verbal rica do PE ultrapassa os anos de estudo formal do português como LNM, o que não se verifica em outras variedades não-europeias do português.

Além disso, o contexto de contato dos timorenses com o português é mais limitado, vide o fato de que, por se reservar à instrução formal e a contextos institucionais, o português ensinado em Timor-Leste se assemelha mais a uma língua estrangeira. Não há, portanto, um *input* robusto que permita maior assimilação das propriedades morfossintáticas do PE, destacando-se, dentre elas, a concordância verbal.

4.2 RESULTADOS GERAIS PARA O SUJEITO NULO

Os resultados a seguir foram obtidos com base no quantitativo de sujeitos plenos pronominais (111) e nulos (111) selecionados do total geral de 582 dados, conforme apresentado na subseção 4.1. Ao realizar o cômputo geral desses dados, a partir da exclusão dos sintagmas nominais da análise, houve uma diminuição expressiva do total de dados, visto que os sujeitos plenos expressos por sintagmas nominais representaram 61% do total geral de dados que compõem o *corpus* (360 de 582).

Além disso, é válido mencionar que alguns textos *apenas* apresentaram sujeitos plenos expressos por sintagma nominal. Isso quer dizer que, dos 80 sujeitos investigados, 9 não produziram, em seus textos, sujeitos nulos ou plenos expressos por pronome.

A seguir, há a apresentação por variáveis dos resultados encontrados para o confronto do sujeito nulo *versus* sujeito pronominal nos dados de escrita do PTL.

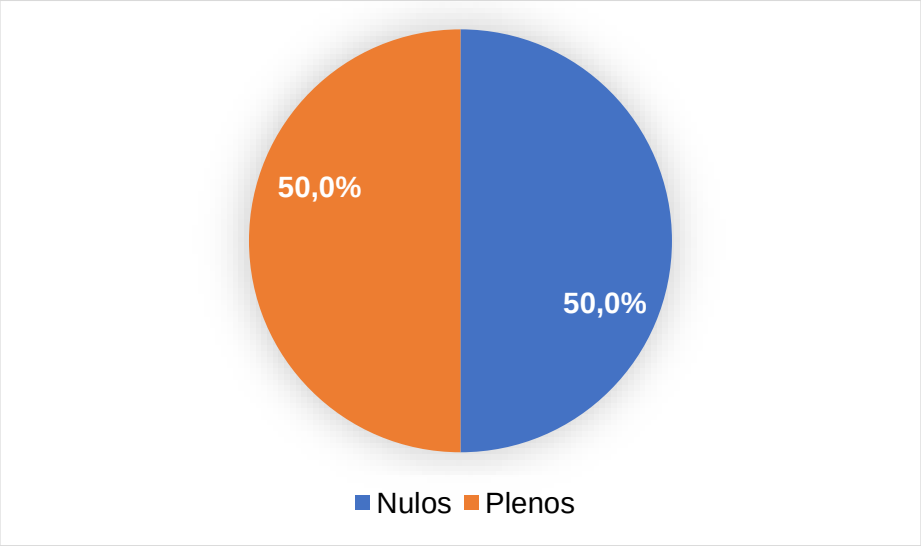
4.2.1 Resultados por variável

4.2.1.1 Tipo de sujeito

Em Timor-Leste, há várias línguas em contexto de multilinguismo. O tétum, língua nacional e reconhecidamente *não-pro-drop* (ALBURQUERQUE, 2011), está presente em toda a escolaridade. Os resultados gerais do processamento dos dados demonstram uma aparente estabilidade na produção de sujeitos nulos e plenos. Do

total de 222 ocorrências, 111 foram com sujeitos nulos e 111 com sujeitos plenos. Em outras palavras, os resultados mostram ocorrências concorrentes de 50% para cada tipo de sujeito, como mostra o gráfico 2.

Gráfico 2 - Distribuição dos sujeitos no PTL (%)



Fonte: autor deste trabalho

A título de constatação, esse resultado geral se aproxima consideravelmente dos percentuais obtidos por Malta (2019) quando analisou textos escritos do PA e do PST. Esses estudos também evidenciaram, como se vê agora, que a produção de sujeitos nulos nessas variedades não é tão expressiva quanto no PE.

a) Sujeito nulo

(175) *pro* gostei muito deste curso (T5, M)
(176) *pro* acho que preciso contribuir as idéias na universidade (T1, M)
(177) *pro* fazemos a relação com multilateral (T26, M)
(178) *pro* Quero continuar na uma universidade (T60, F)

b) Sujeito pleno expresso por pronome

(179) *eles* sempre falaram a lingua portuguesa (T8, M)
(180) *nós* precisamos colaborar (T19, M)
(181) *eles* só adoram pela a natureza (T33, M)
(182) *Eu* gosto de falar e estudar (T66, F)

Adicionalmente, é de grande relevância o fato de que no tétum, língua *não-pro-drop*, o sujeito seja sempre preenchido e isso pode causar nos timorenses a necessidade de preencher essa posição por meio de várias estratégias, de maneira que os pronomes são significativos, gerando essa estabilidade na produção dos sujeitos em relação aos nulos.

4.2.1.2 Morfologia de flexão verbal

Apenas em relação aos 111 sujeitos plenos pronominais analisados, houve maior percentual de concordância padrão, com 77,4%, apontando para uma regra variável. Esse resultado é interessante porque, embora os estudantes tenham concluído o Ensino Médio e estejam submetidos a um contexto de maior monitoramento da língua durante a escrita de suas redações, a morfologia de flexão verbal apresenta-se um pouco enfraquecida:

Quadro 16 - Concordância verbal do sujeito pleno pronominal

Concordância	Ocorrência no *corpus*	%
Padrão	86/111	77,4%
Não padrão	25/111	22,6%

Fonte: autor deste trabalho

a) Concordância padrão

(183) *Eu gostaria* estudar a língua português (T2, M)
(184) *eu queria* expressar o que é que eu aprendo (T44, F)
(185) *nós queremos* falar a lingua portuguesa (T67, F)
(186) *eles sabem* falar a língua portuguesa (T70, F)

b) Concordância não-padrão

(187) *Eu pode* estuda bem (T17, M)
(188) *eu* também *espera* que quando recebeu o bolsa de estudo com UniLab (T24, M)
(189) *nós usa* lingua portuguesa (T49, F)
(190) *ajente podem* falar a lingua portuguêsa (T60, F)

4.2.1.3 Tipo de verbo

Nas orações em que havia sujeitos nulos e plenos expressos por pronome, a maior produção foi de verbos transitivos no geral, alcançando 90,1%. No que diz respeito à produção de sujeitos nulos e plenos, por outro lado, os verbos transitivos apresentaram estabilidade. Já os verbos copulativos foram percentualmente menos produtivos, com 36,4%, e apontam para um maior percentual de uso de sujeitos plenos (63,6%) quando comparados aos verbos transitivos (49,5%) e inacusativos (45,4%):

Quadro 17 - Distribuição de sujeitos nulos e plenos por tipo de verbo

Tipo de verbo	Ocorrência no *corpus*	Sujeitos nulos	Sujeitos plenos
Copulativo	11/222 5%	4/11 36,4%	7/11 63,6%
Transitivo	339/222 90,1%	101/200 50,5%	99/200 49,5%
Inacusativo	11/222 4,9%	6/11 54,6%	5/11 45,4%

Fonte: Autor deste trabalho

a) Copulativos

(190) eu *fico* pronto para servir (T28, M)

(191) eles *são* graduado no tempo da Indonésia (T39, M)

(192) Eu *sou* timorense (T52, F)

(193) *pro esta* localizado no Sudoeste Asiático (T13, M)

b) Transitivos

(194) eu *quero* desenvolver a nosso nação (T11, M)

(195) Eu *gosto* programa Bolsa de estudo (T15, M)

(196) eu *pretendo* continuar o meu studo para conseguir colaboral no desenvolvimento (T18, M)

(197) *pro acho* que *pro preciso* contribuir as idéias na universidade (T1, M)

c) Inacusativos

(198) *voltaram* pra nosso pais (T16, M)

(199) partilhar o que eu *encontrei* durante minha aprendizagem (T34, M)

(200) eu *estudo* ou aprender no curso (T45, F)

(201) eles *vem* para ensinar (T56, F)

Com base nos resultados obtidos, não se percebe nesta variável fator determinante para a realização do sujeito pleno ou nulo, conforme ilustrado no quadro 17.

4.2.1.4 Tipo de oração

Considerando, por um lado, o sujeito nulo e o sujeito pleno pronominal, foram pouquíssimas orações encaixadas: apenas 5 de 222. As orações independentes foram a maioria em todo o *corpus* e isso se demonstrou também nesses dados selecionados.

Para a realização do sujeito nulo, por outro lado, as poucas orações encaixadas foram maioria, 80%. Em relação às outras orações, houve equilíbrio entre sujeitos nulos (49,3%) e plenos (50,7%), conforme apresentado no quadro 18:

Quadro 18 - Distribuição de sujeitos nulos e plenos por tipo de oração

Tipo de oração	Ocorrência no *corpus*	Sujeitos nulos	Sujeitos plenos
Encaixada	5/222 2,3%	4/5 80%	1/5 20%
Outras orações	217/222 97,7%	107/217 49,3%	110/217 50,7%

Fonte: autor deste trabalho

a) Encaixadas

(202) nós estudamos a língua portuguesa *quando pro voltamos para o nosso pais* (T62, F)
(203) Espero *que pode contribuir* (T64, F)
(204) Eu penso *que lingua portugues importante* (T29, M)
(205) *pro acho* que *pro preciso contribuir as idéias na universidade* (T1, M)

b) Outras orações

(206) *pro queria aprender mais* (T9, M)
(207) *Eu fiz aminha candidatura* (T18, M)
(208) *eu pretendo cursar uma universidade* (T48, F)
(209) *pro falamos sobre a lingua portuguesa* (T54, F)

Diante desse resultado, ainda que com poucas orações encaixadas, é possível sinalizar o contexto de orações encaixadas favorecendo o sujeito nulo. No PB, de acordo com Kato e Duarte (2014), as orações encaixadas representam um contexto que favorece o sujeito nulo, sendo uma das características que permitem afirmar que o PB é língua de sujeito nulo parcial. Por outro lado, para as outras orações ocorrem

mais sujeitos plenos. Naturalmente, seria necessário observar num *corpus* mais amplo como se comportam as orações encaixadas no PTL em relação à produção do sujeito nulo, para verificar se essa produção é similar ao que ocorre PB, como esse resultado inicial aparentemente aponta.

4.2.1.5 Pessoa e número do sujeito

Conforme o quadro 19, a seguir, a maioria dos sujeitos nos dados foi de primeira pessoa do singular, 67,9%, com uma leve maior ocorrência de sujeitos pronominais plenos (58,7%). No entanto, com exceção da terceira pessoa do singular, cujas 29 ocorrências foram sujeito nulo, nenhuma outra pessoa e número teve maior quantidade com o sujeito nulo, embora a diferença da liderança do sujeito pleno sempre ficou inferior a 20 pontos percentuais.

Quadro 19 - Distribuição de sujeitos nulos e plenos por referência semântica do sujeito

Pessoa e número	Ocorrência no *corpus*	Sujeitos nulos	Sujeitos plenos
1ª singular	150/222 67,9%	62/150 41,3%	88/150 58,7%
3ª singular	29/222 13,1%	29/29 100%	0/29 0%
1ª plural	25/222 11,3%	12/25 48%	13/25 52%
3ª plural	16/222 7,2%	7/16 43,8%	9/16 56,2%

Fonte: autor deste trabalho

Os resultados acima vão na direção do que está previsto pela Escala de Referencialidade proposta por Kato, Duarte e Cyrino (2000), pois os sujeitos mais referenciais, como os de primeira pessoa, foram mais realizados como sujeitos

pronominais (58,7%). Por outro lado, os sujeitos de terceira pessoa no singular analisados, menos referenciais, foram nulos.

4.2.1.6 Sexo do escrevente

Quanto à variável extralinguística em análise, a produção de sujeitos nulos e plenos entre os timorenses homens e mulheres foi muito parecida, com uma diferença percentual de apenas 3% entre os sexos, conforme mostra o quadro 20:

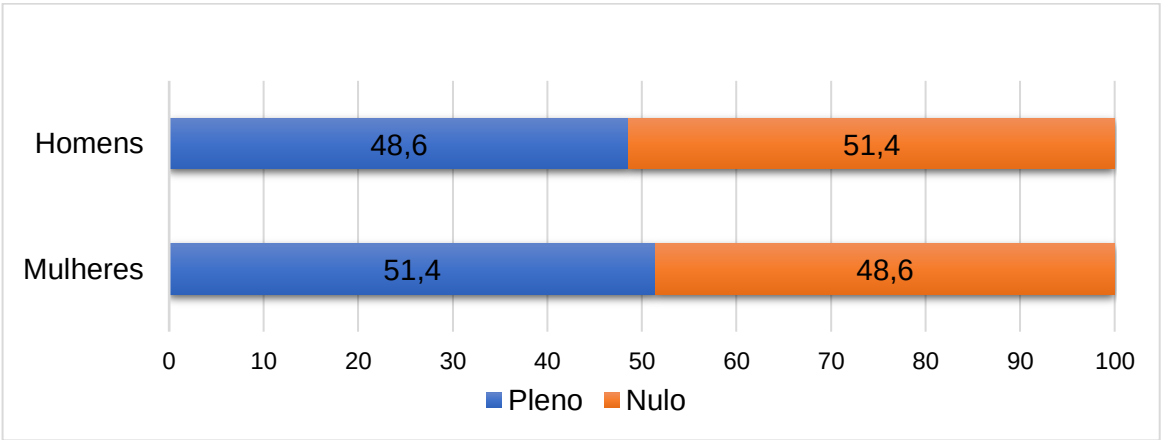
Quadro 20 - Distribuição de sujeitos nulos e plenos de acordo com o sexo do escrevente

Sexo do escrevente	Ocorrência no *corpus*	Sujeitos nulos	Sujeitos plenos
Homens	111/222 50%	57/111 51,4%	54/111 48,6%
Mulheres	111/222 50%	54/111 48,6%	57/111 51,4%

Fonte: autor deste trabalho

Para otimizar a visualização, o gráfico 3 apresenta o par de resultados da produção de sujeitos por homens e mulheres em evidência. Para o sexo masculino, o sujeito nulo representou maioria.

Gráfico 3 - Tipo de sujeito por sexo do escrevente (%)



Fonte: autor deste trabalho

Embora as mulheres tenham produzido 3% a menos de sujeitos nulos que os homens, acredito não ser possível, tomando por base os dados analisados, considerar essa produção demonstrativa de alguma diferença significativa em direção a uma mudança no que diz respeito ao afastamento do PTL em relação à norma europeia do português, que é ensinada em Timor-Leste. Duarte (1995, p. 50), ao analisar essa variável no PB, percebeu que os homens lideraram a produção de sujeitos nulos (46%) em relação às mulheres (32%). A diferença chegou, portanto, a 14% em sua pesquisa, uma situação não verificada nos dados de escrita do PTL.

4.3 DISCUSSÃO DOS RESULTADOS

No conjunto dos fatores apresentados até aqui, foi possível observar como resultados gerais que o PTL, por um lado, apresenta certo distanciamento em relação ao PE no que diz respeito à ausência de diferença quantitativa entre a produção de sujeitos nulos e plenos e à ausência da inversão da ordem S-V. Por outro lado, o PTL também se distancia do PB pela ausência de duplicação do sujeito, pela pouca expressividade de sujeitos de referência não definida e por não apresentar relevante interferência do sexo do escrevente para a produção de sujeitos nulos.

Se o acesso à GU para a aprendizagem do português fosse total (HAT), é bem provável que as características morfossintáticas do PE fossem aprendidas com maior êxito pelos timorenses. No entanto, como esse acesso se dá de maneira parcial (HAP), é possível afirmar que a interferência do multilinguismo causa uma aprendizagem “difícil”, como mencionou Batoreo (2010).

A quantidade de verbos copulativos encontrados na escrita do PTL é interessante, pois a produção de verbos de cópula pelos timorenses é fruto da aprendizagem do PE, uma vez que o tétum, por exemplo, não realiza esses verbos (ALBUQUERQUE, 2011).

Ademais, é importante dar atenção à variável concordância verbal. Conforme já defendido na literatura gerativista, línguas com morfologia de flexão verbal rica permitem a produção de sujeitos nulos, mas, no PTL, com um 77,4% de concordância padrão com sujeitos pronominais, isso não ocorre: sujeitos nulos (50%) possuem o mesmo percentual dos sujeitos pronominais (50%).

O maior percentual de uso da concordância no *corpus* pode estar relacionado com questões de ensino que podem enfocar mais a concordância verbal padrão do que o uso de sujeitos nulos, já que a concordância é um aspecto morfossintático mais saliente na superfície textual, no sentido de ser mais suscetível à estigmatização social do que a manifestação do sujeito nulo. Diniz & Silva (2013) verificam que, entre a população jovem de 15 e 24 anos, os dados oficiais de Timor-Leste em relação a saber ler, falar e escrever uma língua (alfabetização) mostram o seguinte: "A [República Democrática de Timor-Leste] possui 77.8% da população alfabetizada em Tétum, 55.6% em Indonésio, 39.3% em Português, 22.3% em Inglês e 79.1% em qualquer uma dessas quatro línguas."

De fato, o tétum, conforme já tinha mencionado, é a língua majoritária da população. Além disso, nos últimos vinte anos, a quantidade de falantes de português tem aumentado, especialmente por conta da cooperação internacional para o ensino. Albuquerque & Ramos (2020, p. 84) admitem que, em grande parte, o português ensinado em Timor-Leste é o europeu:

> A difusão e o ensino da língua portuguesa em Timor-Leste têm sido objeto de múltiplos projetos e programas de cooperação, em particular originados no Brasil e em Portugal. A cooperação brasileira atuou somente em áreas específicas, com um contingente reduzido e com projetos que não foram renovados ou continuados. Assim, as ações oficiais brasileiras em Timor-Leste apresentaram algumas limitações, conforme será explicado posteriormente. Já a cooperação portuguesa teve um foco maior na difusão e no ensino da língua portuguesa em todas as áreas relacionadas e em todos os níveis de ensino, na política, em diversos setores da sociedade leste-timorense, etc.

Os autores ainda afirmam que um importante programa na cooperação brasileira, comandado pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), o Programa de Qualificação de Docente e Ensino de Língua Portuguesa no Timor-Leste (PQLP), criado em 2005, foi desativado em 2013. Sendo assim, nos mais diversos segmentos, o português tem seu ensino em Timor-Leste realizado, em grande parte, por programas da cooperação portuguesa, que envolvem o governo de Portugal ou universidades portuguesas com o governo e instituições de ensino timorenses.

Dentro desse contexto de programas de cooperação, é possível observar cada vez mais o *status* de Língua Estrangeira que o português tem nesse país atualmente. Outro fator importante é trazido pela pesquisa de Santos (2009), que investigou o

sujeito nulo no PTL. Em sua pesquisa, foram analisados dois grupos distintos: a) um grupo formado por jovens entre 18 e 20 anos que havia estudado português somente na escola, apenas depois da independência de Timor-Leste da Indonésia (grupo 1) e que, segundo Santos (2009, p. 474), aprendeu português como língua estrangeira e b) um grupo formado por adultos de 40 a 55 anos que tinha estudado português antes da ocupação da Indonésia quando Timor ainda era colônia portuguesa e voltou a ter contato com o idioma após a independência de Timor-Leste (grupo 2); esses adultos "já se encontravam num estádio avançado de aquisição da língua." (SANTOS, 2009, p. 474), o que se evidencia pelo maior percentual de uso de sujeitos nulos quando comparados ao uso desses sujeitos pelos jovens, conforme apresentado no quadro 21:

Quadro 21 - Resultados da realização do sujeito em Santos (2009)

	Sujeitos nulos	**Sujeitos plenos**
Grupo 1	39 15%	225 85%
Grupo 2	192 31%	432 69%

Fonte: Santos (2009, p. 475)

Os resultados de Santos (2009) evidenciam que o grupo de jovens que teve contato com o português mais tardiamente e, por isso, aprendeu-o como língua estrangeira, usou mais o sujeito pleno. O sujeito nulo foi expressivamente maior, o dobro, no grupo 2 cujos falantes tinham um contato maior com o português e não o adquiriram como uma língua estrangeira, mas como uma L2, tal como ocorre com o português em Moçambique que, segundo verificou Santos (2010), é falado por 39% da população, convivendo com a L1, que sempre é uma língua da família bantu

Santos (2009) também analisou o uso da concordância verbal pelos dois grupos investigados e chegou aos seguintes resultados:

Quadro 22 - Resultados da concordância verbal em Santos (2009)

	Concordância padrão	Concordância não padrão
Grupo 1	149 57%	114 43%
Grupo 2	584 94%	40 6%

Fonte: Santos, 2009 (p. 475 e 476)

Mais uma vez, os falantes do grupo 1 produzem menos concordância, sendo a regra de concordância verbal variável, tal como verificado nos resultados desta dissertação, o que pode ser fruto da aprendizagem do português como Língua Estrangeira. À semelhança do que foi apresentado na subseção 4.1 em que há uma alto percentual de sintagmas nominais ocupando a posição de sujeito (76,4%), Santos (2009) também verifica que raramente os jovens

> optam por orações com Sujeito Nulo, preferindo a repetição constante do Sujeito para que surja sempre realizado: «Timor-Leste fazer independência no dia 20 Maio. Timor-Leste tem muito história. Timor-Leste tem ocupação portuguesa 250 anos e na ocupação indonésia 24 anos. Timor-Leste tem 13 distritos e 64 suco.» (SANTOS, 2009, p. 476, grifos da autora).

Nesse caso, a autora não relaciona a repetição constante do sujeito como sendo causada por estratégia de coesão. Ao contrário, percebo que, para além de questões textuais e discursivas, o uso constante de sintagma nominais (muitas vezes, repetidos) é característica do PTL e pode ser atribuído à relação ao contato com outras línguas no contexto de multilinguismo, uma análise que pode ser adotada também para esta dissertação. Observem-se trechos de algumas redações que evidenciam o uso constante de sintagmas nominais em posição sujeito:

(210) *Português* é uma língua oficial de Timor-Leste. *Os estudantes* usam no ensino de aprendizagem... Aprender português é muito importante para os timorenses (T5, M)

(211) *Timor-Leste* é uma ilha... *Timor-Leste* consideram como é um asa de costume. *Lingua portuguesa* é muito importante... *Timor-Leste* é uma nação democrática. (T9, M)

(212) *Língua portuguesa* é muito importante. *Timor-leste* tem quatro idioma... *a lingua portuguesa* é muitas facil para nós. *O português* é o idioma que registrar maior crescimento. *A lingua portuguesa* é uma lingua que nós uzamos. (T52, F)

No caso da pesquisa de Santos (2009), o grupo 2 que a autora analisou, além de produzir maior concordância padrão e maior quantidade de sujeitos nulos, também apresentou

> situações de inversão, como se verifica nos exemplos: «Estavam ali alguns macacos.»; «E assim apanhou-o o crocodilo.»; «Apareceu a mãe num repente.» A possibilidade de ocorrência de inversão livre Sujeito-Verbo é outra das propriedades associadas ao valor positivo do Parâmetro do Sujeito Nulo. (SANTOS, 2009, p. 477).

Dessa maneira, a partir dessas importantes considerações e da pesquisa desenvolvida por Santos (2009), é possível perceber a influência da realidade multilíngue no PTL, haja vista que o português é uma língua aprendida em contextos formais como é o caso dos escreventes dos textos desta pesquisa. O tétum, por exemplo, não admite inversão da ordem S-V. Desse modo, os falantes mais jovens não produzem essa inversão.

A uma conclusão similar chegou Martins (2016) ao analisar os erros de aprendizagem de português pelos timorenses falantes de tétum. Ela associa as características morfossintáticas do tétum à dificuldade de internalização do sistema verbal português por parte de alunos timorenses e, por consequência, “erros” são encontrados na apresentação dos verbos nos textos escritos de sua pesquisa. Segundo ela,

> A causa mais evidente prende-se com a interferência linguística do Tétum na aprendizagem da Língua Portuguesa dos aprendentes que participaram neste estudo. Esta interferência explicará tanto os erros cometidos com o uso incorreto do infinitivo, na pessoa e no tempo, como os de concordância sujeito-predicado, tendo em conta que na Língua Tétum os verbos não têm conjugação, usando-se sempre uma forma invariável e sendo os elementos que indicam tempo e pessoa expressos por elementos lexicais autónomos. (MARTINS, 2016, p. 40)

Levando em consideração a maneira como o PTL tem se apresentado nos textos investigados tanto por mim, como por Santos e por Martins, percebo que há certa semelhança entre as análises, no sentido de que as formas que fogem ao que se observa no PE são explicadas via interferência, ou do tétum ou de outras línguas locais, ou, ainda, do contexto multilíngue em geral. Como o PTL atualmente ainda é fruto de aprendizagem, alguns aspectos de sua configuração morfossintática em relação ao PE e a outras variedades do português são diferentes, conforme já referido no início desta subseção.

Aqui, no entanto, distancio-me dos resultados de Santos (2009) no que se refere à sua conclusão. Faço isso porque, metodologicamente, a autora considerou em sua pesquisa o grupo 2 um estágio contínuo do grupo 1 de escreventes. Como são dois grupos que passaram por processos e contextos diferentes em sua relação ao português, não é possível afirmar que, na continuidade da aprendizagem, o grupo 1 apresentará as mesmas características apresentadas pelo grupo 2. Nesse caso, uma pesquisa diacrônica, que acompanhasse o desenvolvimento linguístico dos escreventes em português, dentro de um recorte temporal significativo, seria necessária. Sendo assim, limito-me a concordar com as conclusões da autora sobre a realidade observada para o grupo 1, a saber:

> Nas primeiras fases, o falante realiza as estruturas da LP com base na sua gramática de Tétum, ou seja, sendo o Tétum uma língua de Sujeito obrigatoriamente realizado e com morfologia de concordância verbal pobre, estas propriedades são tidas em conta nas produções em Português. Parece existir, assim, transferência do valor negativo do Parâmetro do Sujeito Nulo. (SANTOS, 2009, p. 477)

Vale a pena ainda assentar que a autora percebeu que, mesmo no estágio mais avançado, o grupo 2 tem

> ainda alguns traços do Tétum, o que nos leva a concluir uma de duas hipóteses: ou estes falantes não estão de facto ainda no estádio final de aprendizagem da L2 ou, por outro lado, atingiram o estádio final de aprendizagem, mas houve alguma fossilização relativamente a estas propriedades. A existência de dificuldades persistentes na realização de propriedades morfológicas é, aliás, uma característica muito frequente do do Tétum nestas fases. (Ibdem, p. 477)

Essa constatação é relevante porque uma vez que, com falantes que têm maior tempo de utilização e maior desenvoltura com a língua portuguesa, persistem características da LM, a conclusão a que chego é a de que a aprendizagem do

português realmente se deu, não pela HAT, como defende a autora, mas pela HAP. Nesse sentido, sou levado a concordar com Albuquerque & Ramos (2020) ao afirmarem que

> O português falado em Timor-Leste é uma variedade nova e específica que ainda está em formação, por isso o caráter de certas estruturas serem variáveis. Assim, os traços linguísticos encontrados, ora considerados erros, peculiaridades ou aspectos da aquisição de L2, são propriedades dessa variedade recente que alguns autores chamam de 'Português de Timor-Leste.' (ALBUQUERQUE; RAMOS, 2020, p. 96).

Sendo assim, considero essencial reconhecer o contexto de desenvolvimento do PTL, bem como a sua aprendizagem pelos timorenses, que se dá por meio da instrução formal dentro de programas estabelecidos para ensino de português ainda organizados pela cooperação internacional. Essas características fazem do PTL uma variedade peculiar do português, cujo *status* é de LNM, mais próximo, evidentemente, de uma língua estrangeira, o que não se observa, por exemplo, no português de Angola, Moçambique e Cabo Verde onde o português é uma L2. Após 20 anos de independência de Timor-Leste e a retomada do ensino de português, a sua reinserção ainda está em curso e a sua morfossintaxe reflete características do contexto multilíngue. Na próxima seção, abordarei um contexto observado no *corpus* que diz respeito a construções com sujeitos em orações infinitivas.

4.4 O CASO DOS SUJEITOS EM ORAÇÕES INFINITIVAS: BREVES INCURSÕES

Como língua com o paradigma de flexão verbal rico, o PE se utiliza de AGR para licenciar e identificar os sujeitos nulos. Sendo assim, é esperado que, quanto mais concordância nos dados, mais sujeitos nulos sejam produzidos. No que se refere à aprendizagem do português por falantes de outras línguas, é possível aventar a hipótese de que, quanto mais exposto for o falante à instrução formal do paradigma de conjugação verbal do português, mais os valores paramétricos relacionados à L2/LNM português serão adequadamente marcados, diferenciando-se dos valores paramétricos da L1/LM. Não obstante, a partir dos dados escritos do PTL analisados nesta dissertação e produzidos por estudantes do Ensino Médio que já estão, portanto,

em um nível de escolaridade avançado, observei ocorrências de sujeitos plenos realizados tanto por pronomes como por sintagmas nominais em orações com verbos não flexionados.

Tanto Santos (2009) como Martins (2016), ao trabalharem em suas pesquisas com textos de timorenses, observaram ocorrências de verbos no infinitivo. Santos não os abordou detalhadamente em sua investigação, relegando as ocorrências a casos de desvio da concordância padrão. Eles representaram, em sua pesquisa, 18% das ocorrências do que ela chamou de “verbos mal conjugados” para o grupo 1 e de 6% para o grupo 2. (SANTOS, 2009, p. 475, 476).

Martins (2016), por sua vez, discute a questão do ponto de vista morfossintático, chamando-os de erros de “uso do infinitivo em vez de presente do indicativo” (MARTINS, 2016, p. 33), optando por “considerar a hipótese de o uso excessivo do infinitivo, em detrimento do uso das outras formas adequadas, implica que o uso destas tem mais dificuldade em ser cabalmente aprendido.” (Ibdem, p. 38)

No que diz respeito aos verbos no infinitivo na presente pesquisa, metodologicamente, essas ocorrências, 33 ao todo, não foram inseridas na análise linguística realizada na seção anterior. Todavia, com o objetivo de entender melhor esse tipo de realização no PTL, submeti esses dados separadamente a uma rodada estatística, cujos resultados apresento a seguir.

Do total de 33 ocorrências, todas ocorrem em orações-raiz. Além disso, todas tiveram o sujeito preenchido, anteposto ao verbo e sem duplicação. Com respeito ao tipo de verbo, em 33 ocorrências, 30 foram transitivos (90,9%), 2 copulativos (6,1%) e 1, inacusativo (3%), conforme apresentado no quadro 23:

Quadro 23 - Tipo de verbo das orações no infinitivo

Tipo de verbo	Total de ocorrências	%
Transitivo	30/33	90,9%
Copulativo	2/33	6,1%
Inacusativo	1/33	3%

Fonte: autor deste trabalho

A respeito da referência semântica do sujeito, eles estão distribuídos entre as primeira e terceira pessoas do singular e do plural, havendo uma maior produção de sujeitos no singular:

Quadro 24 - Referência semântica do sujeito nas orações no infinitivo

Pessoa e número	Total de ocorrências	%
1ª singular	11/33	33,3%
3ª singular	13/33	39,4%
1ª plural	1/33	3%
3ª plural	8/33	24,2%

Fonte: autor deste trabalho

Quanto à variável extralinguística *sexo do escrevente*, houve um equilíbrio nas ocorrências: 17 dados produzidos por homens (51,5%) e 16, por mulheres (48,5%). A seguir, elenco todos os casos de verbos no infinitivo, separados pelo tipo de sujeito pleno, para melhor ilustrar os dados[24]:

a) Orações com verbos no infinitivo com sujeitos expressos por pronomes (12 ocorrências):

(213) eu *voltar* no Brazil para sustentar a minha família (T10, M)
(214) Eu *fazer* o curso (T17, M)
(215) Eu *finalizar* o meu studo (T18, M)
(216) eu já *acabar* a minha estudo na UNILAB (T28, M)
(217) Eu *contribuir* para desenvolver nossas nação (T31, M)
(218) Eu *aprender* o português para implementa (T32, M)
(219) eu *esforçar* muito a minha capacidade (T46, F)

[24] Em adição às ocorrências com o verbo no infinitivo, encontrei em um dos textos a ocorrência a seguir: "a população timorense *decidido* usar essa língua (T64, F)". Nesse caso, usa-se um particípio em vez do verbo flexionado. Acredito ser válido mencionar.

(220) eu pronto *atender* em uma qualquer actividades (T38, F)

(221) nós *comunicar* com lingua portuguesa (T39, F)

(222) eu *aprender* a lingua portuguêsa (T61, F)

(223) eu *atender* a sua nação de Timor-Leste (T61, F)

(224) Eu *voltar* de Brasileiro (T61, F)

b) Orações com verbos no infinitivo com sujeitos expressos por sintagma nominal (21 ocorrências):

(225) o cidadão do Timor-Leste *pegar* um princípio (T1, M)

(226) Timor-Leste *implementar* a lingua português (T7, M)

(227) Timor-Leste com governo do Brazil *criar* um plano (T16, M)

(228) os timorenses *alcançar* a sua independencia (T22, M)

(229) os timorenses *lutar* para sonhar a sua independência (T22, M)

(230) o Timor-Leste *aumentar* o seu plano (T22, M)

(231) a língua portuguesa *ser* como uma língua falada (T27, M)

(232) UNILAB *entregar* o poder para as pessoas (T32, M)

(233) o Timor-Leste *entrar* no CPLP (T34, M)

(234) A lingua portuguesa *ajudar* sobre o processo da aprendizagem (T36, M)

(235) Timor-Leste já *usar* a lingua portuguesa para capacitar mais todos os professores (T39, M)

(236) muitas pessoas *atender* o bolseiro (T48, F)

(237) algumas pessoas de Timor *falar* com a língua Melayu (T51, F)

(238) Timor leste *aprender* a lingua portuguesa (T51, F)

(239) O governo do Timor-leste *dar* as saudações (T52, F)

(240) timorense *utilizar* a lingua portuguesa (T52, F)

(241) O Governo Timor leste *levar* a língua portuguesa (T52, F)

(242) os estudantes sempre *falar* a língua Tétum (T59, F)

(243) os dosentes sempre *usar* a língua melayu (T59, F)

(244) o Timor-Leste *estudar* a lingua português (T61, F)

(245) pessoas timorenses *falar* a língua indonésia (T70, F)

4.4.1 Ocorrência e não ocorrência de infinitivo nos textos do *corpus*

Como é possível observar nas sequências frásicas acima, 21 dos 80 textos apresentaram verbos no infinitivo e eles estão distribuídos entre escreventes homens e mulheres. Interessante é observar que todos os escreventes que produziram formas de verbo no infinitivo também produziram formas finitas. Nos textos que apresentaram verbos no infinitivo, esses verbos não predominaram, mas ocorreu uma, duas e, no máximo, quatro vezes. Por exemplo, o texto 61, que apresentou mais ocorrências de verbos no infinitivo (no total, quatro ocorrências), apresentou também cinco ocorrências de orações com verbos finitos. A título de comparação, observem-se esses exemplos, recolhidos desse texto:

(246) a. A lingua português *é* uma língua oficial (T61, F)
b. Timor-Leste *usam* de lingua portuguêsa (T61, F)

(247) a. eu *aprender* a lingua portuguêsa (T61, F)
b. o Timor-Leste *estudar* a lingua português (T61, F)

Do ponto de vista morfossintático, o escrevente do texto 61 produz períodos simples e usa pouco a concordância padrão. Isso não se revela apenas para uma determinada referência semântica do sujeito. Em (246), temos duas orações, ambas com a referência do sujeito em terceira pessoa do singular. Para o verbo *ser*, foi utilizada a concordância padrão (cf. (246a)). No entanto, para o verbo *usar*, foi empregada a forma do plural em vez do singular, além da inserção de uma preposição cuja regência do verbo não a solicita (cf. (246b)).

Em (247), temos dois casos de verbo no infinitivo. Na primeira oração, o verbo *aprender* está relacionado ao sujeito pronominal *eu* (cf. (247a)). Na segunda, o verbo *estudar* se relaciona ao sujeito de terceira pessoa, *o Timor-Leste* (cf. (247b)). A conclusão a que posso chegar, observando os exemplos deste escrevente, é que há um domínio precário das relações morfossintáticas do português, sendo manifestadas em sua escrita monitorada.

Guasti (2002), abordando a aquisição da linguagem, verificou que construções no infinitivo fazem parte do processo de aquisição no período inicial, quando o valor paramétrico associado ao PSN ainda não foi fixado. Por exemplo, é possível perceber

o uso do infinitivo com sujeitos plenos em posição pré-verbal (cf. (248a) e (248b)) e pós-verbal (cf. (248c)) em línguas como o inglês, o dinamarquês e o francês:

(248) a. Cromer wear glasses. (Eve, 2;0) - inglês
Cromer *usar* óculos
b. Hun sove. (Jens, 2;0) – dinamarquês
she sleep-INF
Ela *dormir*
c. Dormir-INF petit bébé. (Daniel, 1;11) – francês
sleep INF little baby
Dormir pequeno bebê.

(GUASTI, 2002, p. 106; 108-109)

Guasti, ao apresentar esses exemplos, produzidos na fala de crianças entre 1 e 3 anos de idade, argumenta que o que ocorre é que, nesse período de aquisição, há um primeiro estágio de aquisição em todas as línguas: sujeitos ocorrem diante de verbos no infinitivo.

No que se refere à posição do sujeito, em particular, a autora também verifica o uso de sujeitos nulos por crianças entre 2 e 3 anos de idade. Segundo ela, "children learning to speak a variety of languages (e.g., Danish, Dutch, English, French, German) optionally omit the subject of sentences, though their target language requires it to be lexically expressed."[25] (GUASTI, 2002, p. 151). Exemplos disso podem ser vistos em (249).

(249) a. Se, blomster har. (Jens, 2;2) - dinamarquês
look flowers have/has
'Look, (I/you/she/we) have/has flowers.'
Olha, (eu/você/ela/nós) tem/teve flores
b. Tickles me. (Adam, 3;6) – inglês
Cócegas mim
c. Mange du pain. (Grégoire, 2;1) – francês

[25] "Crianças aprendendo a falar uma variedade de idiomas (por exemplo, dinamarquês, holandês, inglês, francês, alemão) opcionalmente omitem o sujeito das frases, embora seu idioma alvo exija que seja expresso lexicalmente." Tradução minha.

eat-3sg some bread
Comer um pouco de pão

(Ibid., p. 151)

Esse fenômeno, exemplificado em (249), chamado de sujeito nulo precoce, é uma demonstração do processo de aquisição: enquanto o valor paramétrico não é fixado, há alternância entre nulos e plenos preenchendo a posição de sujeito. Nesse sentido, conforme abordado por Guasti (2002), a criança produzirá sentenças que não são produzidas pela gramática do adulto. Esse processo natural da aquisição evidencia que os parâmetros da GU estão sendo formatados e essa alternância de produção cessará quando a GU atingir o seu estado estável. No caso da produção de sujeitos nulos, a fixação do PSN ocorre entre os 2 e 3 anos de idade (GUASTI, 2002, p. 185).

Nas ocorrências do PTL, observo que adultos escolarizados, em textos escritos monitorados, produziram, em alternância com verbos finitos, construções com verbos no infinitivo. Isso me leva a afirmar que esses timorenses ainda não fixaram totalmente o valor do PSN da língua-alvo, o português, por interferência da sua gramática internalizada, o tétum – que (1) não admite sujeitos referenciais nulos em orações raízes e (2) não possui um paradigma de flexão verbal rico (DEUS, 2011). Esse resultado, além de expressar dificuldades de aprendizagem do português LNM/LE, também demonstra que o acesso à fixação paramétrica da nova língua em idade adulta passa, evidentemente, pelos parâmetros já fixados na GU durante a aquisição, indo ao encontro do que prevê a Hipótese de Acesso Parcial, que estabelece que, passado o período crítico, qualquer aprendizagem de língua acontecerá com o acesso à GU a partir de sua configuração obtida após a aquisição da LM, condição na qual os valores paramétricos encontram-se já fixados (SCHACHTER,1989; STROZER, 1992).

Desse modo, os dados com verbos no infinitivo são uma evidência adicional no sentido de revelar que o PTL se afasta do PE em relação ao PSN, pois diferentes estágios de desenvolvimento do parâmetro podem ser percebidos nos dados, desde construções infinitivas a construções finitas com sujeitos plenos e nulos. Vale ressaltar ainda que muitos dos escreventes não produziram verbos no infinitivo, mas aqueles que produziram evidenciaram uma alternância entre verbos finitos e infinitos.

Dias (2009), pesquisando o português de Moçambique, observa que, até os falantes fixarem o valor positivo do PSN da língua-alvo (3º estágio), a aquisição desse parâmetro passa por outros dois estágios: o primeiro em que o verbo carrega a morfologia de terceira pessoa do singular, não concordando essa flexão com o traço de pessoa do sujeito, que deve ser pleno, como é o caso de (250a) e o segundo, em que ainda se verifica essa flexão, havendo ocorrência de sujeito nulo e sujeito pleno (cf. (250b)):

(250) a. Eu *inventou* uma palhota (=... inventei...) (1º estágio)
b. Eu *leva* e *foi* vender aquela bola (=... levei...fui...) (2º estágio)
c. *pro* Arranjei uns dois paus. (3º estágio)

(DIAS, 2009, *apud* COMPANHIA, 2018, p. 9)

Em resumo, a autora entende os três seguintes estágios da seguinte forma:

> No primeiro estágio, verifica-se uma tendência para uma fixação negativa deste parâmetro, caracterizada pelo uso do sujeito pleno e pelo enfraquecimento da concordância verbal; no segundo estágio, regista-se uma tendência para uma aquisição progressiva do parâmetro pro-drop com o uso de frases com sujeito pleno e também de frases com sujeito nulo; no terceiro estágio, a interlíngua dos aprendentes caracteriza-se pela fixação positiva do parâmetro pro-drop com o uso de frases com sujeito nulo. (Ibid., p. 9)

Diante das ocorrências apresentadas neste capítulo, é possível observar mais de um estágio de aquisição do PSN nos dados escritos do PTL e isso é verificado dentro de um texto produzido por um mesmo falante. Ou seja, é possível que um mesmo falante produza verbos infinitivos e verbos finitos. Por exemplo, apresento, a seguir, orações produzidas nos textos de número 22 (cf. (251)) e 61 (cf. (252)), escritos por um estudante e uma estudante, respectivamente:

(251) a. os timorenses *alcançar* a sua independência (=...alcançaram...)
b. Timor-Leste *é* nação que *pro alcançar* o mínimo (=...alcançou...)
c. o Timor-Leste e indonesia *são* inimigos

Em (251), o escrevente do sexo masculino não construiu nenhuma oração com sujeito nulo com verbo finito. No entanto, no que se refere à concordância, há variações na sua escrita. Primeiro, o verbo *alcançar* está no infinitivo (cf. 251a)). Há

também, em outro período, o uso do verbo *ser* flexionado e do verbo *alcançar,* na oração encaixada correferente, na forma do infinitivo (cf. (251b)). Além disso, foi produzido o verbo *ser* flexionado adequadamente na terceira pessoa do plural (cf. (251c)).

A já mencionada escrevente do texto 61 também apresentou construções diferentes: a) há sujeito nulo e sujeito pleno com a mesma pessoa do paradigma pronominal (cf. (252b) e (252c)) e b) há sujeito pronominal pleno com verbo no infinitivo (cf. (252a)):

(252) a. eu *aprender* a lingua portuguêsa (=...aprendo...)
b. Eu *vou* de curso de UNILAB
c. *pro* creio conseguir vira a ter recurso nos meus estudos na UNILAB

Diante desses exemplos, observo que há estágios diferentes na aquisição da marcação PSN e do paradigma flexional dos verbos. Como o português é fruto de aprendizagem para esses falantes, pode ser que a LM esteja atuante no surgimento da interlíngua, que seria uma estratégia transitória para se alcançar as características da língua-alvo. Mas, ao mesmo tempo, em vista do quantitativo geral de sujeitos nulos averiguados no *corpus*, bem abaixo do de sujeitos plenos, estaria mesmo o PTL a caminho do sujeito nulo? Para permitir uma reflexão nesse sentido, reconheço que um estudo sincrônico, com um controle dos informantes, acompanhando-os em sua aprendizagem do português ao longo do tempo, seria a medida mais apropriada para aferir que a alternância entre formas finitas e infinitas dos verbos e a escassa produção de sujeitos nulos são apenas reflexo da interlíngua, de modo que essas características tenderiam a desaparecer com o tempo. Ainda assim, esse estudo poderia averiguar, assim como Kato e Duarte (2014) fizeram para o PB, que não há uma mudança em andamento, mas sim uma variedade do português com características próprias.

Pensando dessa maneira, é interessante mencionar que tanto a pesquisa de Santos (2009), com dois grupos distintos em idade e tempo de contato com o português, como a pesquisa de Martins (2016) apresentaram verbos no infinitivo na escrita monitorada dos timorenses em concorrência com verbos em formas finitas. Sendo assim, a assertiva de que o valor associado ao PSN ainda não está fixado pelos timorenses falantes do PTL parece mais acertada, bem como que as línguas já adquiridas têm influência no processo de aprendizagem da língua-alvo.

5 CONSIDERAÇÕES FINAIS

A realidade linguística de Timor-Leste é algo que impressiona. Em um território insular em que vivem 1.167.242 milhões de pessoas[26], várias línguas são faladas. Além das línguas locais, o inglês, o indonésio e o português se fazem presentes, caracterizando a maioria dos habitantes como, ao menos, conhecedor de duas línguas. O tétum desempenha o papel de língua nacional, unindo a comunicação em todo o território e possibilitando a integração do povo. Ao mesmo tempo, o português, historicamente presente, seja diretamente, seja por meio de crioulos, seja na forma de língua estrangeira, continua em ascensão.

As políticas linguísticas em prol da promoção do português estão presentes desde a emancipação política do país. Passados cerca de 20 anos, o seu ensino é difundido por Timor-Leste e vários acordos de cooperação internacional para a ampliação da presença do português no país estão em curso, especialmente com instituições portuguesas.

Do ponto de vista da linguística, a convivência de línguas de *status* diferentes, materna, não materna, estrangeira, promovem um rico fenômeno para a compreensão da Faculdade da Linguagem e a capacidade humana de usar línguas para a comunicação. Nesta pesquisa, a partir da construção teórica em torno do PSN (CHOMSKY, 1981 e seguintes), observei detalhes da língua escrita do português de Timor-Leste. Essa língua, fruto da aprendizagem formal dos timorenses, tem características que a diferencia de outras variedades do português em relação ao PSN.

Quanto aos sistemas de sujeito nulo que podem ser atribuídos ao PE, língua de sujeito nulo consistente, e ao PB, língua de sujeito nulo parcial, não podem ser implementados ainda ao português de Timor-Leste. O motivo é que, de acordo com os dados analisados, há escreventes que ainda não fixaram o valor do PSN: os que usam verbos no infinitivo, por exemplo; no entanto, há indícios de escreventes que fixaram esse valor. Mesmo assim, o PTL não pode ser considerado uma língua de sujeito nulo consistente como o PE. Essa variedade sinaliza uma maior aproximação com o PB, embora se distinga deste em alguns aspectos. Nesse sentido, é plausível supor que o PTL se encontra em uma posição intermediária entre as línguas locais, como o tétum e o PB. Por isso, a interlíngua pode ser verificada nos dados. Diante

[26] De acordo com o Censo Demográfico de 2015 da República Democrática de Timor-Leste.

desse quadro, pesquisas futuras envolvendo a língua falada por timorenses podem ser de grande valia a fim de estabelecer uma comparação com os resultados obtidos neste estudo, bem como de outras variedades do português. Desse modo, poder-se-ia verificar se há distinção entre a língua falada e a língua escrita a respeito do PSN.

Diante do quadro analisado neste trabalho, com a língua escrita, o que se vê é a presença de influências da LM, por exemplo, na ordem de palavras, na quantidade expressiva de sujeitos plenos produzidos no *corpus* e na ocorrência de verbos no infinitivo. Sendo assim, embora exista a expectativa de que o PTL se aproxime mais das características morfossintáticas do PE por ser esta a variedade do português ensinada no país por meio de instrução formal, ainda há uma distância considerável. Isso ajuda a entender que o contexto em que estão os falantes é importante para a investigação do PSN. Em Timor-Leste, por questões históricas e sociais, o português esteve proibido por cerca de 20 anos durante a ocupação indonésia (ALBUQUERQUE, 2011). As consequências disso são perceptíveis, visto que a língua não é materna da população mais jovem. Como resultado, não se aprende português sem a instrução formal, instrução esta que é ministrada em tétum (BATOREO, 2010; MARTINS, 2016).

Ademais, os dados analisados evidenciam que o PTL e o PE não compartilham o mesmo *status* em relação à concordância e ao PSN: no primeiro caso, a regra de concordância verbal é variável no PTL e semicategórica no PE e, no segundo caso, esta produz mais sujeitos nulos do que aquela. Nesse sentido, é possível aventar a seguinte hipótese: o que os estudantes timorenses de fato aprendem é o português com uma gramática em formação.

Neste estudo, portanto, foi possível caracterizar o comportamento da concordância e do PSN em dados escritos do PTL, valendo ressaltar que essa análise não esgota, de maneira alguma, a compreensão do tema. É fato que a fascinante capacidade humana da linguagem merece cada vez mais atenta investigação e acredito ser o PTL mais uma excelente fonte para ampliação e compreensão de como construímos o conhecimento linguístico internalizado.

REFERÊNCIAS

ABNEY, S. P. *The English Noun Phrase in its Sentential Aspect.* Tese (Doutorado em Filosofia) – MIT, Cambridge, 1987.

ALBUQUERQUE, D. B. *Esboço gramatical do Tetun Prasa: língua oficial de Timor-Leste.* Dissertação (Mestrado em Linguística) - Universidade de Brasília, Brasília, 2011.

ALBUQUERQUE, D. B. Esboço morfossintático do português falado em Timor-Leste. *Moderna Spraak,* v. 106, n. 1, p. 1-10, 2012.

ALBUQUERQUE, D. B; RAMOS, R. O português na República Democrática de Timor-Leste: quase duas décadas depois. *Revista Letras Raras*. Campina Grande, v. 9, n. 4, p. 76-105, dez. 2020.

ALEXIADOU, A; ANAGNOSTOPOULOU, E. Parametrizing Agr: word order, verb-movement and EPP checking. *Natural Language and Linguistic Theory*, 16:3, 1998, p. 491-539.

ALMEIDA, N. C. H. *Língua Portuguesa em Timor-Leste: Ensino e Cidadania*. Dissertação (Mestrado em Língua e Cultura Portuguesa). Universidade de Lisboa, Lisboa, 2008.

ANTUNES, I. C. *Lutar com palavras: coesão e coerência*. São Paulo: Parábola Editorial, 2005.

BAKER, M. *Thematic roles and syntactic structure.* In: HAEGEMAN, L. (Ed.) *Elements of Grammar.* Dordrecht: Kluwer, 1996.

BARBOSA, P. *Null Subjects*. Tese (Doutorado) - MIT, Cambridge, 1995.

BARBOSA, P. Two kinds of subject pro. *Studia Linguistica,* 63-1, 2009, pp. 2-58.

BATORÉO, H. J. Ensinar português no enquadramento poliglóssico de Timor-Leste. *Palavras*, n. 37, Lisboa: Associação de Professores de Português, 2010, p. 55-65.

BRANDÃO, S. F; VIEIRA, S. R. Concordância nominal e verbal: contribuições para o debate sobre o estatuto da variação em três variedades urbanas do Português. *Alfa: Revista de Linguística* (UNESP. Online). v. 56 (3), 2012, p. 1035-1064.

BRANDÃO, S. F; VIEIRA, S. R. Tipologia de regras linguísticas e Estatuto das variedades/línguas: a concordância em português. *Lingüística*. Vol. 30 (2), Diciembre 2014, p. 81-112

BRANDÃO, S. F; VIEIRA, S. R. The agreement continuum in urban samples of African, Brazilian and European varieties of Portuguese. In: LÓPEZ, l. A.; GONÇALVES, P; AVELAR. J. O. (Org.). *The Portuguese Language Continuum in Africa and Brazil*. Amsterdam/the Netherlands: John Benjamins Publishing Co, 2018, p. 267-289.

BRITO, R. P. Temas para a compreensão do atual quadro linguístico de Timor-Leste. *Ciências & Letras*, Porto Alegre, n. 48, p. 175-194, jul./dez. 2010.

BUNGE, M. *La ciencia, su método y su filosofia.* Buenos Aires: Siglo Veinte, 1974.

CARDOSO, H. C.; HAGEMEIJER, T.; ALEXANDRE, N. Crioulos de base lexical portuguesa*. Anthologies, textes, attestations et sources des langues romanes*, p. 670-692, 2015.

CARIOCA, C. R. As funções sociais da língua e as políticas de difusão do Português no Timor-Leste. *DELTA.* v. 32, n. 2, p. 427-447, 2016.

CINQUE, G. A null theory of phase and compond stress. *Linguistic Inquiry*, v. 24, n. 2, 1993, p. 239-297.

CHOMSKY, N. *Aspects of the Theory of Syntax.* Cambridge: MIT Press, 1965.

CHOMSKY, N. *Linguagem e pensamento*. Petrópolis: Vozes, 1971.

CHOMSKY, N. Conditions on transformations. In: ANDERSON, S.R; KIPARSKY, P. *A festschrift for Morris Halle*. Holt, New York: Rinehart and Winston, 1973.

CHOMSKY, N. *Reflections on Language.* Pantheon Books, 1975.

CHOMSKY, N. On Wh-Movement. In: CULICOVER, P; WASOW, T; AKMAJIAN, A. (eds). *Formal Syntax*. New York: Academic Press, 1977.

CHOMSKY, N. *Reflexões sobre a linguagem*. São Paulo: Cultrix, 1980.

CHOMSKY, N. *Lectures on Goverment and Bilding.* Dordrecht: Foris, 1981.

CHOMSKY, N. *Some Concepts and Consequences of the Theory of Government and Binding*. Cambridge: MIT Press, 1982.

CHOMSKY, N. *Knowledge of Language*: Its Nature, Origin, and Use*.* New York: Praeger, 1986.

CHOMSKY, N. Some notes on economy of derivation and representation. In: FREIDEN, R. (Ed.) *Principle and parameters in comparative grammar*. Cambridge: MIT Press, 1991. p. 417-454.

CHOMSKY, N. A minimalist program for linguistic theory. In: HALE, K.; KEYSER, S. J. (Eds.). The view from building 20 : essays in linguistics in honor of Sylvain Bromberger. *Current studies in linguistics*, 24, Cambridge: The MIT Press, 1993.

CHOMSKY, N. *Linguagem e mente*. Brasília: Editora Universidade de Brasília, 1998.

CHOMSKY, N. *Language and other cognitive systems*. What is special about language?. *Language learning and development*, v. 7, n. 4, p. 263-278, 2011.

COMPANHIA, C. A. Panorama geral dos estudos sobre a aquisição do Português L2 em Moçambique. *Odisseia*, v. 3, n. 1, p. 1-15, jan.-jun., 2018.

COSTA, L. Línguas de Timor. In: CRISTÓVÃO, F. *Dicionário Temático da Lusofonia*. Lisboa: Texto Editores, 2005.

CYRINO, S. M. L.; DUARTE, M. E. L.; 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.* Madrid: Iberoamericana, 2000, p. 55-73.

DEUS, A. S. *O ensino da língua portuguesa em Timor-Leste: o método Português em Timor e a importância do Tétum (L1) na aquisição do português (L2).* Dissertação (Mestrado em Ensino do Português como L2 e LE) Universidade Nova de Lisboa, Lisboa, 2009.

DEUS, A. S. O Tétum-Díli como Língua Não-Pro-Drop: Na Senda do Caboverdiano. *Anais do XXVI Encontro Nacional da Associação Portuguesa de Linguística*, FLUP: Porto, 2011, p. 226-241.

DIAS, H. O parâmetro do sujeito nulo em aprendentes do português em Moçambique. *Português moçambicano: Estudos e reflexões*, p. 17-46, 2009.

DUARTE, M. E. L. *A perda do princípio 'Evite Pronome' no português brasileiro.* Tese (Doutorado em Linguística) - Unicamp, Campinas, 1995.

DUARTE, M. E. L. The loss of the 'avoid pronoum' principle in Brazilian Portuguese. In: KATO, M. A.; NEGRÃO, E. V. (Eds.). *Brazilian Portuguese and the null subject parameter.* Madrid: Iberoamericana, 2000, p. 17-36.

DUARTE, M. E. L. A evolução na representação do sujeito pronominal em dois tempos. In: Paiva, M. C. & Duarte, M. E. L Mudança Linguística em Tempo Real. *Contra Capa*, p. 115-128, 2003.

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

FARACO, C. A. *História sociopolítica da língua portuguesa*. São Paulo: Parábola Editorial, 2016.

FERRARI-NETO, J. As bases e os objetivos da Gramática Gerativa. In: FERRARI-NETO, J.; SILVA, C. R. T. (org.) *Programa minimalista em foco: princípios e debates*. Curitiba: CRV, 2012.

FINGER, I. Aquisição de segunda língua: abrangência e limitações do modelo gerativista. *Revista da ABRALIN*, v. 2, n. 2, 2003, p. 31-57.

FREIRE. O. Sobre o parâmetro do sujeito nulo: interferência do português brasileiro no espanhol falado por nativos residentes no Brasil. Dissertação (mestrado) - UFPE, Centro de Artes e Comunicação, Programa de Pós-graduação em Letras, 2013.

FODOR, J. A. *The Modularity of Mind*. Cambridge: MIT Press, 1983.

GALVES, C. C. *Ensaios sobre as gramáticas do português.* Campinas: Ed. da Unicamp, 2001.

GALVES, C. et al. *Parameter theory and linguistic change.* Oxford: Oxford University Press, 2012.

GONÇALVES, M. F. H. S. L. *Para uma redefinição do parâmetro do sujeito nulo*. 1994. Dissertação (Mestrado em Linguística Portuguesa Descritiva) – Faculdade de Letras da Universidade de Lisboa, Lisboa, 1994.

GONÇALVES, P. Pesquisa linguística e ensino do Português L2: Potencialidades das taxonomias de erros. *Linguística: Revista de Estudos Linguísticos da Universidade do Porto*, v. 2, p. 61-76, 2007.

GUASTI, M. T. *Language acquisition*. Cambridge: MIT Press, 2002.

GUIMARÃES, J. E. Língua portuguesa em Timor-Leste: análise de orientações para o ensino da escrita. *PERcursos Linguísticos*, v. 7, n. 14, p. 503-522, 2017.

HAUSER, M. D.; CHOMSKY, N.; FITCH, W. T. The faculty of language: What is it, who has it, and how does it evolve? *Science*, 298, 2002, p. 1569–1579.

HOLMBERG, A. Is There a Little Pro? Evidence from Finnish. *Linguistic Inquiry,* v. 36, 2005, pp. 533-564.

HOLMBERG, A. Null subject parameters. In: BIBERAUER, T; HOLMBERG, A; ROBERTS, I; SHEEHAN, M. *Parametric variation: Null subjects in minimalist theory*. Cambridge: Cambridge University Press, 2010

HOLMBERG, A.; ROBERTS, I.; Introduction: parameters in minimalist theory. In: BIBERAUER, T. et al. *Parametric variation*: Null subjects in minimalist theory. Cambridge University Press, 2009.

HOLMBERG, A.; NAYUDU, A; SHEEHAN, M. Three partial null subject languages: a comparison of Brazilian Portuguese, Finnish and Marathi. *Studia Linguistica*, v. 63(1), p. 59-97, 2009.

HUANG, C.T. J. *Logical Relations in Chinese and the Theory of Grammar*. Tese (Doutorado em Filosofia) – MIT, Cambridge, 1982.

HUANG, C. T. J. On the Distribution and Reference of Empty Pronouns. *Linguistic Inquiry*, 15:4, 1984, pp. 531-574.

HUANG, C. T. J. Remarks on empty categories in Chinese. *Linguistic inquiry*, 1987, p. 321-337.

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

HULL, G. A Morphological Overview of the Timoric Sprachbund. *Studies in Languages and Cultures of East Timor*, v.4, p.98-205, 2001b.

JACKENDOFF, R. *X-Syntax: a Study of Phrase Structure.* Cambridge: MIT Press, 1977.

KAIL, M. *Aquisição da linguagem*. São Paulo: Parábola Editorial, 2013.

KATO, M. A. The partial pro-drop nature and the restricted VS order in Brazilian Portuguese. In: KATO, M. A.; NEGRÃO, E. V (Eds.). *Brazilian Portuguese and the null subject parameter*. Madrid: Vervuert, Iberoamericana, 2000. p. 223-258.

KATO, M. A. Sintaxe e aquisição na Teoria de Princípios e Parâmetros. *Letras de hoje*, Porto Alegre, v. 30, n. 4, dezembro 1995, p. 57-73.

KAISER, G. A. Sobre a (alegada) perda do sujeito nulo no português brasileiro. In: LOBO, T. (ed.) et al. *Novos dados, novas análises*; Tomo 1. Salvador: Editora da Universidade Federal da Bahia, p. 11-42, 2006.

KENEDY, E. *Curso básico de linguística gerativa.* São Paulo: Contexto, 2013.

KOOPMAN, H.; SPORTICHE, D. The position of subjects. *Língua*, v. 85, n. 2, p. 211-258, 1991.

KRZEK. M. The null subject parameter meets te Polish impersonal -NO/-TO construction. In: *Order and structure in syntax II*: Subjecthood and argument structure. Berlin: Language Science Press, p. 307–318, 2017.

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

LABOV, W. Some sociolinguistic principles. In. PAULSTON, C. B; TUCKER, G. R. (eds.) *Sociolinguistics: the essential readings*. Massachusetts: Blackwell Publishing, 2003.

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

LASNIK, H.; SAITO, M. *Move alpha: conditions on its application and output*. Cambridge: MIT Press, 1992.

LIGHTFOOT, D. *How to set parameters: arguments from language change*. Cambridge: MIT Press, 1991.

LOPES, F. J.; CAMPOS, E. A. A expressão pronominal no português brasileiro e no português falado em Cabo Verde - trilhando possíveis (as)simetrias. *PAPIA-Revista Brasileira de Estudos do Contato Linguístico*, v. 25, n. 2, p. 319-345, 2015.

MADEIRA, A. Aquisição de L2. In: OSÓRIO, P; GROSSO, M. J*. Português Língua Segunda e Língua Estrangeira – Da(s) Teoria(s) à(s) Prática(s)*. Lisboa: Lidel, 2004.

MAGALHÃES, T. M. V. *Aprendendo o sujeito nulo na escola.* Dissertação (mestrado) - Universidade Estadual de Campinas, Instituto de Estudos da Linguagem, Campinas, SP, 2000.

MALTA, D. P. L. N. *A natureza da morfologia de flexão verbal e o parâmetro do sujeito nulo em dados escritos de variedades africanas do português: uma análise contrastiva*. Dissertação (Mestrado em Linguística) UFPE, Recife, 2019.

MARCELINO, M. Aquisição de segunda língua e bilinguismo. *Revista Intercâmbio*, v. XXXV. P. 38-67, 2017.

MARTINS, B. *Interferências Linguísticas na Aprendizagem do Português por Falantes do Tétum.* Dissertação (Mestrado em Português Língua Não Materna (PLNM) – Português Língua Estrangeira (PLE) / Língua Segunda (L2)) – Universidade do Minho, Braga, 2016.

MATEUS, M. H. M. et al. *Gramática da língua portuguesa*. v. 6. Lisboa: Caminho, 2003.

MAY, R. *Logical form: Its structure and derivation*. Cambridge: MIT Press, 1985.

MIOTO, C. A gramática gerativa e a aquisição da linguagem, *Letras de Hoje*, 30, pp. 75-82, 1995.

MIOTO, C; SILVA, M. C. F.; LOPES, R. *Novo Manual de Sintaxe*. São Paulo: Contexto, 2018.

MODESTO, M. Sujeitos nulos em línguas de tópico proeminente. *Revista da ABRALIN*, v. 3, n. 1, p. 119-145, 2004.

MOGFORD, K.; BISHOP, D. Desenvolvimento da linguagem em condições normais. In: MOGFORD, K.; BISHOP, D. (orgs). *Desenvolvimento da linguagem em circunstâncias excepcionais.* Rio de Janeiro: Revinter, p. 1-26, 2002.

MOTA, M. A; MIGUEL, M; MENDES, A. A concordância de P6 em português falado. Os traços pronominais e os traços de concordância. *Revista Papia*, v. 22, n. 1, p. 161-187, 2012.

MUFWENE, S. S. *The Ecology of Language Evolution*. Cambridge: Cambridge Press, 2001.

NEWMEYER, F. J. Against a parameter-setting approach to typological variation. *Linguistic Variation Yearbook*, v. 4, n. 1, 2004, p. 181-234.

NICOLIS, M. The null subject parameter and correlating properties. In: BIBERAUER T. (Ed). *The limits of syntactic variation*, John Benjamin Publishing, v. 132, 2008, p. 271-294.

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 JUNIOR, R. *Aspectos morfossintáticos do português brasileiro: o sujeito nulo em textos escritos.* 2018. Trabalho de Conclusão de Curso (Licenciatura em Letras - Português e Espanhol) UFRPE, Recife, 2018.

PARAÍSO, T. C. F. *O uso da concordância nominal de número em redações produzidas em português por estudantes guineenses e timorenses: um estudo contrastivo*. Dissertação (Mestrado em Linguística) UFPE, Recife, 2019.

PARKVALL, M. *Da África para o Atlântico*. Campinas: Editora da Unicamp, 2012.

PEREIRA, D. *Crioulos de base portuguesa*. Lisboa: Caminho, 2007.

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

PLATZACK, C. Agreement and the Person Phrase hypothesis. *Working Papers in Scandinavian Syntax*, 73, 2004, p. 83-112.

POLLOCK, J. *Langage et Cognition*. Paris: Presses Universitaires de France, 1997.

PUPP SPINASSÉ, K. Os conceitos Língua Materna, Segunda Língua e Língua Estrangeira e os falantes de línguas alóctones minoritárias no Sul do Brasil. *Revista Contingentia*, vol. 1, p. 01-10, nov 2006.

RANDRIAMASIMANANA, C. *Word order typology and Malayo-Polynesian languages*. Massey University, 2019.

RAPOSO, E. P. *Teoria da gramática*: a faculdade da linguagem. Lisboa: Caminho, 1992.

RIZZI, L. *Issues in Italian Syntax*. Dordrecht: Foris, 1982.

RIZZI, L. Negation, Wh-Movement and the Null Subject Parameter, 1982. In: KAYNE, R.; LEU, T.; ZANUTTINI, R. *An Annotated Syntax Reader*: Lasting Insights and Questions. Oxford: Wiley-Blackwell, 2014.

RIZZI, L. Two notes on the linguistic interpretation of Broca's aphasia. In: *Agrammatism*. Academic Press, 1985, p. 153-164.

RIZZI, L. Null subjects in Italian and the theory of pro. *Linguistic Inquiry*, v. 17, n. 3, p. 501-558, 1986.

RIZZI, L. The new comparative syntax: principles and parameters of universal grammar. In: 10th ADVANCED COURSE “LANGUAGE AND COGNITION”. Foundation Archives Jean Piaget, Geneva, p. 1-20, October 10, 1988.

RIZZI, L. A parametric approach to comparative syntax: properties of the pronominal system. HAEGEMAN, L. (Ed.). *The New Comparative Syntax*. London, New York: Longman, 1997. p. 268-285.

SANKOFF, D; TAGLIAMONTE, S; SMITH, E. *Goldvarb X: A variable rule application for Macintosh and Windows*. Department of Linguistics, University of Toronto, 2005.

SANTOS, A. S. R. A influência da L1 no processo de aquisição da L2: um estudo sobre a transferência de parâmetros morfológicos e sintácticos. Textos Seleccionados. *XXIV Encontro Nacional da Associação Portuguesa de Linguística*, Lisboa, APL, 2009, pp. 467-481.

SANTOS, A. M. B. O português moçambicano entre o português brasileiro e o português europeu. *Filologia e Linguística Portuguesa*, v. 12, n. 1, p. 27-39, 2010.

SCHACHTER, J. Testing a proposed universal. In: S. GASS; J. SCHACHTER (Eds.). *Linguistic perspectives on second language acquisition*. Cambridge: Cambridge University Press, 1989.

SCHERRE, M. M. P; NARO, A. J. Análise quantitativa e tópicos de interpretação do Varbrul. In: MOLLICA, M. C.; BRAGA, M. L. (Orgs.). *Introdução à Sociolinguística: o tratamento da variação*. São Paulo: Editora Contexto, 2012, p. 147-177.

SLABAKOVA, R. *Second language acquisition*. Oxford: Oxford Core Linguistics, 2016.

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*. Tese (Doutorado em Linguística) - UFAL, Maceió, 2004.

SOUZA, C. M. N. et al. O preenchimento do sujeito pronominal em textos escritos de alunos adolescentes de Florianópolis. *Working Papers em Linguística*, v. esp., p. 94-107, 2010.

STOWELL, T. *Origins of phrase structure.* 1981. Dissertação de PhD – MIT.

STROZER, J. Non-native language acquisition from a principles and parameters perspective. In: CAMPOS, H; MARTNEZ-GIL, F (Eds.). *Current studies in Spanish linguistics*. Georgetown: Georgetown University Press, 1992.

TARALDSEN, K. T. *On the nominative island constraint, vacuous application and the that-trace filter*. Bloomington: Indiana University Linguistics Club, 1978.

VERCEZE, R. M. A. N. GERATIVISMO: SUAS CONTRIBUIÇÕES PARA A LINGUÍSTICA. *Revista Philologus*, v. 15, n. 43, 2009, p. 93-100.

VERÍSSIMO, V. A evolução do conceito de parâmetro do sujeito nulo. *Entrepalavras*, Fortaleza, v. 7, jan./jun. 2017, p. 76-90.

WILLIAMS, E. On the notions 'Lexically Related' and 'Head of a Word'. *Linguistic Inquiry,* 12, 1981, p. 245-274.