



UNIVERSIDADE FEDERAL DA BAHIA
INSTITUTO DE LETRAS
PROGRAMA DE PÓS-GRADUAÇÃO EM LÍNGUA E CULTURA

YANNA KAROLINA FIGUEIREDO DE SOUZA

**O DESENVOLVIMENTO DO SISTEMA DE DETERMINANTE DE REFERÊNCIA
DEFINIDA NO PORTUGUÊS DE MONTE CAFÉ (SÃO TOMÉ)**

Salvador

2021

YANNA KAROLINA FIGUEIREDO DE SOUZA

**O DESENVOLVIMENTO DO SISTEMA DE DETERMINANTE DE REFERÊNCIA
DEFINIDA NO PORTUGUÊS DE MONTE CAFÉ (SÃO TOMÉ)**

Dissertação apresentada ao programa de Pós-Graduação em Língua e Cultura, como parte dos requisitos para obtenção do título de Mestre em Letras.

Salvador

2020

Sistema de Bibliotecas - UFBA

Souza, Yanna Karolina Figueiredo de.

O desenvolvimento do sistema de determinante de referência definida no português de Monte Café (São Tomé) / Yanna Karolina Figueiredo de Souza. - 2021.

102 f.: il.

Orientador: Prof. Dr. Alan Norman Baxter.

Dissertação (mestrado) - Universidade Federal da Bahia, Instituto de Letras, Salvador, 2021.

1. Sociolinguística. 2. Linguagem e línguas - Variação. 3. Língua portuguesa - Estudo e ensino - Monte Café (São Tomé, Ilha de (São Tomé e Príncipe)). 4. Língua portuguesa - Artigo - Monte Café (São Tomé, Ilha de (São Tomé e Príncipe)). I. Baxter, Alan Norman. II. Universidade Federal da Bahia. Instituto de Letras. III. Título

CDD - 469.799

CDU - 81'27(669.95)

UNIVERSIDADE FEDERAL DA BAHIA
INSTITUTO DE LETRAS
PROGRAMA DE PÓS-GRADUAÇÃO EM LÍNGUA E CULTURA

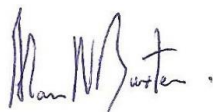
YANNA KAROLINA FIGUEIREDO DE SOUZA

**O DESENVOLVIMENTO DO SISTEMA DE DETERMINANTE DE REFERÊNCIA
DEFINIDA NO PORTUGUÊS DE MONTE CAFÉ (SÃO TOMÉ)**

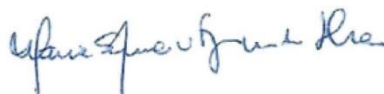
Dissertação apresentada como requisito para obtenção do grau de Mestre em Letras, ao Programa de Pós-Graduação em Língua e Cultura da Universidade Federal da Bahia.

Salvador, janeiro de 2021

Banca Examinadora



Orientador
Prof. Dr. Alan Norman Baxter
Universidade Federal da Bahia (UFBA - Presidente)



Prof^a. Dr^a. Maria Cristina Vieira de Figueiredo Silva
Universidade Federal da Bahia (UFBA – Examinadora Interna)



Prof^a. Dr^a. Shirley Freitas Sousa
Universidade da Integração Lusófona Afro-Brasileira (UNILAB – Examinadora Externa)

A Monte Café.

AGRADECIMENTOS

Essa dissertação representa a finalização de um processo de três anos que, embora árduos, trouxeram mais aprendizados do que eu poderia prever. Às pessoas com quem me conectei nesse caminho, toda a minha gratidão.

Agradeço à “minha torre”, que vejo de onde quer que eu vá no mundo: a minha irmã, Alyce, por despertar o melhor que há em mim a cada dia; a minha mãe, Yoná, por ser essa fonte inesgotável de amor e zelo; a meu pai, Kleber, por ser esse alicerce incansável. À Família Figueirêdo, por me inspirar com tanta determinação. À Família Souza, por me mostrar tamanha resiliência. A Cláudia, Toni, Zenóbia e Chico, pelo vínculo sincero.

Agradeço ao meu orientador, Alan Norman Baxter, por confiar em mim e me orientar com tanta atenção e precisão, mesmo estando do outro lado do mundo. Por cada incentivo, oportunidade e suporte oferecidos incansavelmente, desde 2014, obrigada. Agradeço a Silvana Ribeiro pela orientação rigorosa e cuidadosa durante o Estágio Docente, e aos meus alunos de Secretariado Executivo, por tornarem essa experiência incrível. Aos professores Maria Cristina Figueiredo, Paloma Moore, João Paulo Cyrino, Carla Ferreira, Tânia Lobo, Juliana Gayer e Cláudia Sobrinho, pela prestatividade constante. Agradeço às colegas Larissa Santana e Luiza Menezes, por todas as trocas nesse caminho. A Gregory Guy e Marie-Eve Bouchard, pelas conversas academicamente descontraídas.

A Rebeca Moreira, Gabriela Araújo, Victor Montalvão, Rodrigo Mota, Artur Mota, Patrick Willis e Eric Barretto, pelo acolhimento inesgotável. Aos olhos atentos e mãos dispostas a construir junto: Anna Carolina, Jéssica Farias, Welber Trindade, João Daniel e Caio Vidreira, obrigada pelas revisões e traduções. A Marthinha Böker e Paulo Tiano, por não me permitirem me distanciar da arte. Agradeço aos amigos do Instituto de Letras, do CCAA, da Argentina, do CEVIC, do Gota D'água, do São José e do Análise, pela torcida incessante.

Agradeço, por fim, à Fundação de Amparo à Pesquisa do Estado da Bahia (FAPESB) pelo financiamento desta pesquisa.

“É que pó sabé isso, nosso aqui nosso lingua já é diferente co ôtro.”

A.

RESUMO

Esta dissertação investiga a variação no uso dos determinantes de referência [+ definida] [+específica] no sintagma determinante (SD) em posição de sujeito e objeto direto, no português da comunidade tonga de Monte Café, em São Tomé (África). Essa comunidade compartilha uma série de características com o Português Afro-Brasileiro (LUCCHESI, BAXTER, RIBEIRO, 2009; NEVES, 2014, 2019), sociolinguísticas e históricas: colonização portuguesa, distribuições latifundiárias, exploração de populações do continente africano e situações de contato linguístico e bilinguismo. A dissertação estuda os fatores linguísticos e extralinguísticos que condicionam a variação no uso de determinantes de referência definida, que manifesta três variantes: (i) a realização do artigo definido, como em “A minha geração, são poucos que falam *a língua* dos pais”; (ii) o demonstrativo não-dêitico, como no exemplo “Os patrões é que falavam isso, que criô *essa separação* de nossos avós”; (iii) o determinante zero (\emptyset), como em “*Guera* veio assi memo, nossos preto é que ranjô isso”. A pesquisa focou no uso do Det \emptyset , a variante mais frequente, e se desenvolveu com o aporte teórico-metodológico da Sociolinguística Variacionista (LABOV, 2008; GUY; ZILLES, 2007). Dados extraídos de um *corpus* de 18 entrevistas sociolinguísticas estratificadas em três faixas etárias, 1 (de 20 a 40 anos), 2 (de 41 a 60 anos), e 3 (≥ 61 anos), foram tratados pelo programa estatístico GOLDFARB-X (SANKOFF; TAGLIAMONTE; SMITH, 2005). A análise avaliou a influência de dez variáveis independentes, e identificou como estatisticamente significativas quatro variáveis linguísticas (número no SD, presença de outro constituinte no SD, animacidade do substantivo, e familiaridade da referência) e duas variáveis sociais (faixa etária e sexo). Confirmou-se a relação estreita ente SD singular e o uso do Det \emptyset , como também o efeito favorecedor da presença de outros constituintes no SD. A variável familiaridade da referência revela o papel da pragmática no uso do Det \emptyset , enquanto a animacidade do substantivo confirma a relação proposta por Lyons (1999) entre os traços [+ definido] e [+ animado] e o determinante nulo. Os resultados realçam a importância da interface entre o nível gramatical e o da pragmática para derivar interpretações corretas da leitura [+ específica] [+ definida] nesta variedade de português. No âmbito extralinguístico, as faixas etárias revelam um perfil aquisicional, com uso mais conservador pelos falantes das faixas 2 e 3 (ambas favorecendo o Det \emptyset), e um uso mais inovador entre os falantes da faixa 1 (que favorecem o artigo foneticamente realizado). Por outro lado, a variável sexo mostra um comportamento mais conservador na fala das mulheres (que realizam mais o Det \emptyset) em relação à dos homens.

Key words: Português Tonga; SDs nus; determinante definido; aquisição do artigo definido; Sociolinguística.

ABSTRACT

This dissertation investigates variation in the use of [+ definite] [+ specific] reference determiners in subject and direct object determiner phrases (SD), in the Portuguese of the Tonga community of Monte Café, in São Tomé (Africa). This community shares a number of characteristics with Afro-Brazilian Portuguese (LUCCHESI, BAXTER, RIBEIRO, 2009; NEVES, 2014, 2019), both sociolinguistic and historical: Portuguese colonization, large estates, exploitation of continental African populations and situations of linguistic contact and bilingualism. The dissertation studies the linguistic and extralinguistic factors that condition the variation in the use of definite reference determiners, which displays three variants: (i) the realization of the defined article, as in “A minha geração, são poucos que falam *a língua* dos pais”; (ii) the non-deictic demonstrative, as in “Os padrões é que falavam isso, que criou *essa separação* de nossos avós”; and, (iii) the null determiner (\emptyset), as in “*Guera* veio assi memo, nossos preto é que ranjô isso”. Focusing on the use of the null determiner (Det \emptyset), the most frequent variant, the research was developed within the theoretical-methodological framework of Variationist Sociolinguistics (LABOV, 2008; GUY; ZILLES, 2007). Data extracted from a corpus of 18 sociolinguistic interviews stratified in three age groups, 1 (from 20 to 40 years old), 2 (from 41 to 60 years old), and 3 (\geq 61 years old) were processed by the GOLDVARB-X (SANKOFF; TAGLIAMONTE; SMITH, 2005) statistical program. The analysis assessed the influence of ten independent variables, identifying four linguistic variables (number in the SD, presence of another constituent in the SD, animacy of the noun, and familiarity of the reference) and two social variables (age and gender) as statistically significant. The close relationship between the singular SD and the use of Det \emptyset was confirmed, as well as the favorable effect of the presence of other constituents in the SD. The familiarity of reference variable revealed the role of pragmatics in the use of Det \emptyset , while the animacy of the noun confirmed the relationship proposed by Lyons (1999) between the features [+ defined] and [+ animated] and the null determiner. The results highlight the importance of the interface between the grammatical and pragmatic levels, in order to derive correct interpretations of the [+ specific] [+ definite] reading in this variety of Portuguese. Where extralinguistic variables are concerned, the age groups reveal an acquisitional profile, with more conservative use by the speakers of age-levels 2 and 3 (both favoring Det \emptyset), and a more innovative use among the speakers of age-level 1 (who favor the phonetically realized determiner). On the other hand, the gender variable displays a more conservative behavior in the speech of the women (who use Det \emptyset more) in relation to that of men.

Key words: Tonga Portuguese; bare noun; definite determiner; acquisition of definite article; Sociolinguistics.

LISTA DE FIGURAS

Mapa 1 - São Tomé e Príncipe	58
---	-----------

LISTA DE GRÁFICOS

Gráfico 1 - Perfil global da variável dependente na fala da comunidade de Monte Café	69
Gráfico 2 - Perfil global da frequência da variável dependente binária na fala da comunidade de Monte Café	71
Gráfico 3 - Efeito (P.R.) do número no SD sobre a realização do determinante \emptyset	73
Gráfico 4 - Distribuição percentual da realização do determinante de acordo com o número no SD	73
Gráfico 5 - Efeito (P.R.) do número morfológico do SD sobre a realização do determinante \emptyset .	74
Gráfico 6 - Distribuição percentual da realização do determinante de acordo com o número morfológico no SD	75
Gráfico 7 - Efeito (P.R.) da presença de outros constituintes no SD sobre a realização do determinante \emptyset	77
Gráfico 8 - Distribuição percentual da realização do determinante de acordo com outros constituintes no SD.....	78
Gráfico 9 - Efeito (P.R.) da faixa etária sobre a realização do determinante \emptyset	79
Gráfico 10 - Distribuição percentual da realização do determinante por faixa etária	80
Gráfico 11 - Distribuição percentual da realização do determinante por faixa etária em variedades do português	81
Gráfico 12 - Efeito (P.R.) do sexo do falante sobre a realização do determinante \emptyset	82
Gráfico 13 - Distribuição percentual da realização do determinante por sexo.....	82
Gráfico 14 - Efeito (P.R.) da familiaridade sobre a realização do determinante \emptyset	84
Gráfico 15 - Distribuição percentual da realização do determinante de acordo com a familiaridade.....	85
Gráfico 16 - Efeito (P.R.) da animacidade do substantivo sobre a realização do determinante \emptyset	86
Gráfico 17 - Distribuição percentual da realização do determinante de acordo com a animacidade do substantivo.....	87
Gráfico 18 - Efeito (P.R.) da função sintática do SD sobre a realização do determinante \emptyset	89
Gráfico 19 - Distribuição percentual da realização do determinante de acordo com a função sintática do SD.....	90

Gráfico 20 - Efeito (P.R.) da animacidade do substantivo sobre a realização do determinante Ø	91
Gráfico 21 - Distribuição percentual da realização do determinante de acordo com o tipo de posse	92
Gráfico 22 - Efeito (P.R.) da posição do SD na oração sobre a realização do determinante Ø ...	94
Gráfico 23 - Distribuição percentual da realização do determinante de acordo com a posição do SD	95

LISTA DE TABELAS

Tabela 1 - Efeito do número no SD sobre a realização do determinante Ø.....	73
Tabela 2 - Efeito do número morfológico no SD sobre a realização do determinante Ø.....	75
Tabela 3 - Efeito da presença de outros constituintes no SD sobre a realização do determinante Ø	77
Tabela 4 - Efeito da faixa etária sobre a realização do determinante Ø	79
Tabela 5 - Efeito do sexo sobre a realização do determinante Ø.....	82
Tabela 6 - Efeito da familiaridade sobre a realização do determinante Ø.....	84
Tabela 7 - Efeito da animacidade do substantivo sobre a realização do determinante Ø.....	86
Tabela 8 - Efeito da função sintática do SD sobre a realização do determinante Ø.....	89
Tabela 9 - Efeito do tipo de posse sobre a realização do determinante Ø.....	92
Tabela 10 - Efeito da posição do SD sobre a realização do determinante Ø.....	94

LISTA DE QUADROS

Quadro 1 - Principais fontes dos DLP dos informantes de Monte Café	60
Quadro 2 - Variáveis independentes linguísticas	62
Quadro 3 - Variáveis independentes sociais	66
Quadro 4 - Frequência das variantes da variável dependente na fala da comunidade de Monte Café	Erro! Indicador não definido.
Quadro 5 - Frequência das variantes da variável dependente na fala da comunidade de Monte Café	68
Quadro 6 - Frequência da variável dependente na fala da comunidade de Monte Café	71
Quadro 7 - Variáveis independentes selecionadas e rejeitadas pelo GOLDVARB-X.....	71

LISTA DE ABREVIATURAS E SÍMBOLOS

[±arg]	-	±Argumento
[±pred]	-	±Predicativo
A	-	Adjetivo
Adv	-	Advérbio
AdvP	-	Adverbial Phrase/ Sintagma Adverbial
ART	-	Artigo
AP	-	Adjective Phrase/ Sintagma Adjetivo
CCV	-	Crioulo Caboverdiano
Det Ø	-	Determinante zero
D	-	Determinante
DEM	-	Demonstrativo
DET	-	Determinante
DLP	-	Dados Linguísticos Primários
L1	-	Primeira Língua (ou Língua Materna)
L2	-	Segunda Língua
MC	-	Monte Café
N	-	Nome
NP	-	Noun Phrase/ Sintagma Nominal
Num	-	Numeral
NumP	-	Numeral Phrase/ Sintagma Numeral
NURC	-	Projeto da Norma Urbana Linguística Culta
OD	-	Objeto Direto
OI	-	Objeto Indireto
P	-	Preposição
PTG	-	Português dos Tongas
PA	-	Português de Almojarife
PB	-	Português Brasileiro
PABH	-	Português Afro-Brasileiro de Helvécia
PMC	-	Português de Monte Café

PMN	-	Parâmetro do Mapeamento Nominal
PP	-	Prepositional Phrase/ Sintagma Preposicional
SD	-	Sintagma Determinante
SN	-	Sintagma Nominal
ST	-	Santome (ou Crioulo de São Tomé)
V	-	Verbo
VP	-	Verbal Phrase/Sintagma Verbal

SUMÁRIO

1. INTRODUÇÃO	20
2. OS SINTAGMAS DETERMINANTES (SDs): DAS TEORIAS AOS USOS	23
2.1 LYONS (1999): O CONCEITO DE DEFINITUDE E NOÇÕES CORRELACIONADAS	23
2.1.1 Tipos de familiaridade	23
2.1.1.1 Uso situacional	24
2.1.1.2 Conhecimento geral.....	25
2.1.1.3 Uso anafórico.....	25
2.1.1.4 Referência cruzada	26
2.1.1.5 Uso catafórico.....	26
2.1.2 Especificidade	27
2.1.3 Genericidade	28
2.1.4 Contabilidade e número	30
2.1.5 Animacidade	32
2.1.4 Tipos de posse	33
2.2 UMA PERSPECTIVA SEMÂNTICO-SINTÁTICA DOS SDs NO PORTUGUÊS.....	34
2.2.1 A Teoria X-Barra e a Hipótese do SD	34
2.2.2 O Parâmetro do Mapeamento Nominal de Chierchia (1988)	35
2.2.3 A insuficiência do Parâmetro do Mapeamento Nominal para classificar o PB e a proposta de Schmitt e Munn (1999)	36
2.3. PESQUISAS SOBRE SDs NUS DE REFERÊNCIA DEFINIDA EM VARIEDADES DO PORTUGUÊS.....	39
2.3.1 Os SDs nus [+ específicos][+ definidos] no Português Brasileiro (WALL, 2013) 39	
2.3.2 Os SDs nus no Português Afro-Brasileiro de Helvécia	41

2.3.2.1 A variação na realização do artigo definido nos SNs [+ definidos] de Helvécia (BAXTER; LOPES, 2009).....	41
2.3.2.2 O uso dos SDs nus objetos no PABH: um olhar direcionado para os dados referenciais (NEVES, 2019).....	44
2.3.3 Os SDs nus no Português de Almojarife (FIGUEIREDO, 2019)	47
2.4. PESQUISAS RECENTES SOBRE SDs NUS DE REFERÊNCIA DEFINIDA EM LÍNGUAS CRIOULAS DE SÃO TOMÉ.....	49
2.4.1 Os SDs nus no Crioulo Caboverdiano (BAPTISTA, 2007).....	50
2.4.2 Os SDs nus no Santome (ALEXANDRE E HAGEMELJER, 2007).....	52
2.5 SÍNTESE	55
3. A COMUNIDADE, OS DADOS E A METODOLOGIA	56
3.1 A AQUISIÇÃO DO PORTUGUÊS NA GRANDE ROÇA DE MONTE CAFÉ.....	57
3.2 A CONSTITUIÇÃO DO <i>CORPUS</i> TONGA.....	59
3.3 A VARIÁVEL DEPENDENTE E AS VARIÁVEIS INDEPENDENTES	61
3.3.1. A variável dependente	61
3.3.2 As variáveis independentes linguísticas.....	62
3.3.3 As variáveis independentes sociais.....	65
3.4 O TRATAMENTO DOS DADOS	66
3.5 SÍNTESE	67
4. RESULTADOS.....	68
4.1 A DISTRIBUIÇÃO GERAL DE SDs REFERENCIAIS DEFINIDOS NAS POSIÇÕES DE SUJEITO E OBJETO.....	68
4.2 PERFIL GLOBAL DA VARIÁVEL BINÁRIA E RESULTADOS DAS ANÁLISES: GRUPOS SELECIONADOS PELO GOLDVARB-X.....	70
4.2 GRUPOS SELECIONADOS	72
4.2.1. Número do SD.....	72

4.2.2 Presença de outro constituinte no SD	76
4.2.3 Faixa etária	79
4.2.4 Sexo	81
4.2.5 Familiaridade	83
4.2.6 Animacidade do substantivo	86
4.3 GRUPOS NÃO SELECIONADOS PELO GOLDVARB-X	88
4.3.1. Função sintática do SD	88
4.3.2 Tipo de posse	91
4.3.3 Posição do SD na oração	93
4.4 SÍNTESE	95
5. CONSIDERAÇÕES FINAIS	97
REFERÊNCIAS	99

1. INTRODUÇÃO

Esta dissertação é desenvolvida no âmbito das investigações sobre os Sintagmas Determinantes (SDs) nus, neste caso, do Português da comunidade de Monte Café, em São Tomé (África).

A partir da observação de que os SNs nus possuem restrições em algumas línguas e ocorrem livremente em outras, Chierchia (1998) propõe que as projeções máximas nuas podem ser predicativas (denotam propriedades) e/ou argumentais (denotam diretamente entidades). A partir dos traços [\pm arg] e [\pm pred], essa classificação tipológica é sistematizada no Parâmetro do Mapeamento Nominal (PMN). Assim, segundo Chierchia (1998), há línguas do tipo [+arg, -pred], como o chinês; línguas [+arg, +pred], como o inglês e a maior parte das línguas germânicas; e línguas [-arg, +pred], em que se enquadram as línguas românicas. Infere-se, então, que o português deveria ser uma língua do tipo [-arg, +pred]. Schmitt e Munn (1999), porém, mostram evidências de que o PB possui algumas características atribuídas aos três tipos de línguas, mas não se encaixa em nenhum deles. Como os SNs nus singulares não possuem especificação semântica de número nem leitura de massa no PB, os autores propõem que os SNs nus singulares sejam interpretados como SDs com determinantes vazios e sem número.

Investigando os SDs singulares nus definidos em dados do Português Afro-Brasileiro de Helvécia (Bahia), Baxter e Lopes (2009) afirmam que a formação de um sistema de referências que inclui a ausência do determinante definido pode estar ligada a questões de economia de sistema, de acordo com o grau de previsão de referência atribuído pelos contextos linguísticos e pragmáticos.

O fenômeno investigado na presente dissertação é a variação no uso do determinante em sintagmas determinantes [+ específicos][+ definidos], em posição de sujeito e de objeto direto, a partir de dados de informantes pertencentes às três primeiras gerações a utilizar o português-L1 na comunidade de Monte Café, onde houve um contexto peculiar de aquisição linguística através do contato.

A grande roça de Monte Café é uma grande empresa agrícola, produtora de cacau e café, fundada em 1854 por colonos portugueses. A mão de obra contratada para trabalhar nessa comunidade veio sobretudo de Angola, e os descendentes desses angolanos, nascidos em São Tomé, são denominados tongas. Em Monte Café, os tongas adquiriram o português em uma

situação de Transmissão Linguística Irregular (BAXTER, 1995; LUCCHESI; BAXTER, 2009) em um contexto bastante diverso, devido às variedades de português L1 e L2 às quais tiveram acesso. O regime contratual contínuo levava para a roça, a cada 3-5 anos, novos trabalhadores contratados, crescendo, portanto, o número de adultos falantes de um estágio inicial de português-L2 (BAXTER, 2002, 2004). Considerando a complexidade desse contexto sociolinguístico, é muito provável que no português dos tongas de Monte Café (doravante PMC) haja vestígios de fases iniciais da aquisição do português como L2, que foram incorporados na gramática L1 da comunidade. Salienta-se, assim, a importância de estudar a variação no uso dos determinantes nos SDs sujeitos e objetos diretos [+ específicos][+ definidos] do PMC e seus condicionantes, no intuito de entender como desenvolveu o sistema de marcação de definitude e referencialidade nessa variedade do português.

Esta dissertação tem como principais objetivos: (i) investigar a variação no uso do determinante em sintagmas determinantes sujeitos e objetos [+ específicos][+ definido] nas faixas etárias 1 (20-40 anos), 2 (41-60 anos) e 3 (> 60 anos); (ii) investigar os fatores linguísticos e extralinguísticos que condicionam o uso das variantes do determinante definido: o determinante zero, como em “*Guera* veio assi memo, nossos preto é que ranjô isso”, e os determinantes fonéticos (o artigo definido, como no exemplo “A minha geração, são poucos que falam *a língua dos pais*”, e o demonstrativo, como em “Os patrões é que falavam isso, que criô *essa separação de nossos avós*”); e (iii) observar o desenvolvimento determinante no SD [+ específico][+ definido] no PMC, a partir do tempo aparente. Para isso, a variação no uso dos determinantes foi investigada a partir dos dados de 18 entrevistas sociolinguísticas com os tongas de Monte Café, gravadas em 1998 (BAXTER, 1998-2000). Os informantes apresentam perfil social semelhante quanto à escolaridade (analfabetos ou semianalfabetos) e têm contato mínimo com variedades do português faladas fora da roça, embora os homens tenham mais acesso à variedade padrão devido ao contato com os colonos portugueses.

O trabalho é organizado da seguinte forma: no capítulo 2, é introduzida a fundamentação teórica do estudo, tendo em vista as particularidades semânticas e sintáticas dos SDs reportadas na literatura científica. O capítulo é construído a partir de textos considerados especialmente relevantes dentro da temática: Lyons (1999), Abney (1987) e Chierchia (1998). Em seguida, busca-se observar como o fenômeno se comporta em variedades do português, inicialmente de forma mais teórica, a partir de Schmitt e Munn (1999), e posteriormente em estudos quantitativos do

Português Brasileiro (PB) (WALL, 2013), do Português Afro-Brasileiro de Helvécia (PABH) (BAXTER; LOPES, 2009; NEVES, 2019) e do Português de Almojarife (PA) (FIGUEIREDO, 2019). Inclui-se, ainda, uma breve discussão de pesquisas sobre duas línguas crioulas faladas em São Tomé, o Crioulo Caboverdiano e o Santome, cujos nomes nus possuem características semelhantes tanto às variedades mencionadas como ao Português de Monte Café.

O capítulo 3 apresenta o contexto histórico que teria possibilitado um processo de aquisição particular, em situação de contato, na grande roça de Monte Café, o perfil social dos informantes e as suas principais fontes de dados linguísticos primários (DLP). Apresenta-se, ainda, a variável dependente, que encerra o artigo definido, o demonstrativo e o determinante zero, bem com as variáveis independentes que instrumentalizam hipóteses sobre a realização do determinante, formuladas com base na fundamentação teórica do capítulo 2. A codificação dos dados, conforme essas hipóteses, e a sua quantificação são realizadas sob a perspectiva da Sociolinguística laboviana por meio do pacote estatístico Goldvarb-X.

Os resultados das análises são apresentados e comentados no capítulo 4, considerando os fundamentos teóricos explorados nos capítulos anteriores e o contexto sócio-histórico da comunidade de Monte Café.

2. OS SINTAGMAS DETERMINANTES (SDs): DAS TEORIAS AOS USOS

Neste capítulo, apresenta-se a revisão de literatura sobre os Sintagmas Determinantes (SDs) nus. Uma vez que o objeto de estudo é composto pelos SDs nus de referência [+ específica][+ definida], torna-se imprescindível que as suas características semânticas e sintáticas sejam observadas. Na seção 2.1, o trabalho de Lyons (1999) conceitua a *definitude* e a associa a outras noções às quais está estreitamente interligada. A seção 2.2 aborda a estrutura dos SDs a partir da Hipótese do SD de Abney (1987), a classificação das línguas naturais através do Parâmetro do Mapeamento Nominal de Chierchia (1998) e a proposta de Schmitt e Munn (1999) para classificar os SNs nus singulares no PB. Em seguida, a seção 2.3 e 2.4 apresentam, respectivamente, pesquisas sobre SDs nus de referência definida em variedades do português e em línguas crioulas faladas em São Tomé. Por fim, a seção 2.5 sintetiza o que foi exposto.

2.1 LYONS (1999): O CONCEITO DE DEFINITUDE E NOÇÕES CORRELACIONADAS

Lyons (1999, p. 1-2) constata que em muitas línguas há um elemento cuja função principal é indicar a definitude ou indefinitude dos sintagmas nominais (SNs): o *artigo*. Complementa, porém, que nem todos os SNs contêm um artigo, e a distinção de (in)definitude não se limita apenas às frases que o contêm. Ele afirma que, provavelmente em todas as línguas, a diferenciação entre os SNs definidos e indefinidos também pode ser realizada por meio de outros componentes, como o pronome indefinido e o demonstrativo¹.

O autor explica a noção de definitude, relacionando-a a outras que estão intrinsecamente ligadas à primeira e são, por isso, basilares para o seu entendimento. Nas próximas seções, apresentaremos resumidamente as que se mostraram relevantes também para a análise dos dados deste trabalho, começando pela **familiaridade** e prosseguindo com as noções de especificidade, genericidade, contabilidade e número, animacidade e posse.

2.1.1 Tipos de familiaridade

¹ Lyons (1999, p. 2) explica que, em inglês, a leitura do SN *this house* “esta casa” seria considerada definida e a de *several houses* “várias casas”, indefinida.

A partir da perspectiva da consciência do falante e do ouvinte, Lyons (1999, p.2) apresenta a **familiaridade**: um SN é familiar quando sua referência é conhecida por ambos. Essa familiaridade é, geralmente, evidenciada pelo uso do artigo definido (p. 3). Para elucidar esse conceito são utilizadas as frases (1) e (2)², nas quais os dois *carros* são referências particulares e de conhecimento do enunciador, mas apenas na segunda frase a familiaridade é perceptivelmente compartilhada com o receptor. A **hipótese da familiaridade**, apresentada por Christophersen (1939)³, fundamenta essa diferença: o uso do artigo definido *the* ‘o/a/os/as’ evidencia que o referente é familiar para os interlocutores, enquanto o *a* ‘um/uma’ é utilizado quando não se mostra que o conhecimento do referente é compartilhado (LYONS, 1999, p. 3).

1) I bought *a car* this morning.

Eu comprei *um carro* nessa manhã.

2) I bought *the car* this morning.

Eu comprei *o carro* nessa manhã.

A fim de ilustrar os diversos usos através dos quais um SN pode se tornar familiar para o receptor, eles serão brevemente apresentados e exemplificados⁴ – uso situacional, anafórico, referência cruzada, conhecimento geral e uso catafórico, respectivamente (LYONS, 1999, p. 3).

2.1.1.1 Uso situacional

O uso situacional ocorre quando o falante e o ouvinte se encontram na mesma situação física, o que contribui para que o referente do SN definido se torne familiar (LYONS, 1999, p.4). Os exemplos a seguir ilustram esse emprego:

3) Just give *the shelf* a quick wipe, will you, before I put this vase on it.

Você poderia dar uma limpada rápida *na prateleira* antes de eu colocar o vaso nela.

² Lyons (1999, p. 2). A tradução de todos os exemplos que são originalmente em inglês é nossa.

³ apud Lyons (1999, p. 3).

⁴ Inicialmente, ilustramos os diferentes usos com os exemplos trazidos por Lyons (1999) em inglês (todos com tradução nossa). Em seguida, para cada tipo de uso, foram elaborados dois exemplos no português brasileiro (PB), nos quais os SNs definidos ocupam as funções sintáticas de sujeito e de objeto direto, visto que essas são as funções dos SNs analisados nesta pesquisa.

4) Put these clean towels in *the bathroom* please.

Coloque essas toalhas limpas *no banheiro*, por favor.

5) *O quarto* está limpo, mas *a cama* ainda está desarrumada, está vendo?

6) Aqui está *o livro*, pode levar.

2.1.1.2 Conhecimento geral

Nos exemplos (7), (8), (9) e (10), a familiaridade ocorre através do **conhecimento geral**. Nesses casos, o repertório sociocultural geral permite que *a lua*, *o universo* e *o papa* sejam termos utilizados para se referir a entidades únicas, tendo como referência o mundo em que estamos situados (LYONS, 1999, p. 4). No caso do exemplo (7), Lyons (1999) explica que, ainda que o receptor não conheça pessoalmente *o presidente de Gana*, ele é familiar por sua existência ser de conhecimento público.

7) *The president of Ghana* is visiting tomorrow.

O presidente de Gana vai visitar amanhã.

8) *The moon* was very bright last night.

A lua estava muito brilhante ontem à noite.

9) *O universo* está em constante expansão.

10) Eu vi *o papa* na televisão ontem.

2.1.1.3 Uso anafórico

Os SNs definidos das frases (11), (12) e (13) tornam-se familiares para o receptor através do uso **anafórico**, i.e., já havia sido feita uma referência esse SN no contexto linguístico em que se encontra (LYONS, 1999, p. 4).

11) An elegant, dark-haired woman, a well-dressed man with dark glasses, and two children entered the compartment. I immediately recognized *the woman*. *The children* also looked vaguely familiar.

Uma mulher elegante de cabelo escuro, um homem bem vestido de óculos escuros e duas crianças entraram na cabine. Eu imediatamente reconheci *a mulher*. *As crianças* também pareceram vagamente familiares.

12) Tenho uma história para te contar, mas não pode ser agora. *A história* é longa.

13) Entrei numa livraria ontem, peguei um livro para ler e acabei deixando *o livro* em uma prateleira aleatória.

2.1.1.4 Referência cruzada

Os exemplos (14) e (15), bem como o (16) e o (17), demonstram o que se denomina **referência cruzada** ou **uso associativo**, que é como uma combinação do uso anafórico com o conhecimento geral (LYONS, 1999, p. 4). Espera-se que a menção de um SN como *um táxi*, em (14), acione, a partir do repertório sociocultural do ouvinte, todas as coisas relacionadas a *táxi*, como *banco*, *motorista*, *tarifa etc.* Devido à familiaridade prevista, esses elementos podem ser referidos através de SNs definidos:

14) I had to get a taxi from the station. On the way *the driver* told me there was a bus strike.

Eu tive que pegar um táxi na estação. No caminho, *o motorista* me disse que havia uma greve de ônibus.

15) They've just got in from New York. *The plane* was five hours late.

Eles acabaram de chegar de Nova York. *O avião* estava 5 horas atrasado.

16) Pensei que estava atrasada, mas quando entrei no avião, *o piloto* ainda estava do lado de fora.

17) Ela entrou no carro para dirigir e, depois de se sentar, bateu *a porta* com força.

2.1.1.5 Uso catafórico

Nas sentenças (18), (19), (20) e (21) encontra-se o uso **catafórico**, i.e., o referente se torna familiar para o receptor a partir de uma explicação subsequente ao SN (LYONS, 1999, p. 5), como as orações relativas, em (18), (20) e (21), e o complemento nominal, em (19):

18) *The bloke* Ann went out last night phoned a minute ago.

O cara com quem Ann saiu na noite passada telefonou há um minuto.

19) *The fact* that you've known them for years is no excuse.

O fato de que você os conhece há anos não é desculpa.

20) *A menina* que eu conheci no show me convidou para sair.

21) Ontem eu comprei *o livro* do qual eu te falei na semana passada.

As noções de definitude e de familiaridade estão, portanto, intrinsecamente relacionadas; quando os SNs são definidos, a identificação do seu referente é explicitamente compartilhada pelos interlocutores (LYONS, 1999, p. 6). Outra noção semântico-pragmática apresentada por Lyons (1999) que contribui para o entendimento da definitude é a de **especificidade**, apresentada a seguir.

2.1.2 Especificidade

Para tratar da **especificidade**, Lyons (1999, p. 165) parte inicialmente de dois exemplos desconhecidos para o receptor, portanto, indefinidos. O que os difere é que o primeiro se refere a uma entidade particular, i.e., o enunciador é capaz de identificar o referente, que é, por isso, [+ específico], como *um carro* em (22). Se o referente não é particularizado por quem o menciona, ele corresponde um membro arbitrário de uma classe descrita pelo SN, sendo, então, [- específico], como *um livro* em (23). Apesar de essa ser uma discussão feita majoritariamente em relação aos SNs indefinidos, Lyons (1999, p. 165) afirma que um debate semelhante pode ocorrer em relação aos definidos.

22) I bought *a car*.

Comprei *um carro*.

23) Pass me *a book*.

Me passe *um livro*.

No que concerne à definitude, essa noção é complementar à de especificidade: enquanto a definitude trata da identificação feita pelo ouvinte, a especificidade trata da identificação do referente pelo falante. A fim de sistematizar a relação entre essas noções, as ilustraremos com

exemplos elaborados a seguir. Na sentença (24), o SN destacado é [+ específico][+ definido], pois os interlocutores conhecem *a professora* em questão. Em (25), o emissor sabe quem é *o cantor*, mas o destinatário não, tratando de um SN [+ específico][- definido]. O SN do exemplo seguinte (26) não pode ser identificado pelo falante no momento da enunciação, sendo, por isso, [- específico], mas, a partir do conhecimento geral, sabe-se que alguém ganhou o campeonato. *O vencedor* é de conhecimento do receptor, que esteve presente no jogo e o anunciará em seguida, sendo, a partir disso, conhecido por ambos, [+ definido]. A sentença (27), por sua vez, possui um SN cujo referente que equivale a qualquer membro da categoria *ovo de páscoa*, sendo, assim, [- específico] [- definido].

24) Vi *a professora* dos nossos filhos na rua hoje. [+ específico][+ definido]

25) Descobri *um cantor* incrível no fim de semana. [+ específico][- definido]

26) Eu trabalhei hoje à tarde e não pude ir ao estádio com você. Quem foi *o vencedor* do jogo?
[- específico][+ definido]

27) Ganho *um ovo de páscoa* todo ano na Semana Santa. [- específico][- definido]

Os SNs [- específico] [- definido] estão relacionados à generalização, que será discutida a seguir a partir da noção de **genericidade**. Optou-se por incluir a explicação na seção seguinte deste trabalho, visto que o entendimento de conceitos como genericidade e indefinidade foram necessários para delimitar mais precisamente o *corpus* desta pesquisa, restrito aos SNs [+ específico] [+ definido]. Ademais, diversos estudos sobre SNs nus⁵ mostram a genericidade como fator relevante para que essa estrutura ocorra nas línguas.

2.1.3 Genericidade

A noção de **genericidade** está nos SNs que universalizam uma classe por inteiro e todas as entidades que a ela pertencem (LYONS, 1999. p. 179), em oposição aos nomes [+ específicos]. Lyons (1999, p. 181), evidencia que, dentre as formas utilizadas em inglês para expressar generalização, normalmente não são permitidas frases definidas plurais⁶. No PB, porém, os SNs

⁵ Schmitt e Munn (1999), Müller (2002), Alexandre e Hagemeyer (2007), Baptista (2007).

⁶ Lyons (1999, p. 181) explica que o uso de frases definidas plurais para fazer uma referência genérica é possível com “apenas alguns tipos de nomes, como nomes de nacionalidades, e alguns nomes que denotam classes de classes”. Tradução nossa.

definidos plurais também podem funcionar de forma generalizante, como na tradução do exemplo (31), ao se referir aos cachorros em geral.

28) *A dog* has four legs.

Um cachorro tem quatro patas.⁷

29) *The dog* has four legs.

O cachorro tem quatro patas.

30) *Dogs* have four legs.

Cachorros têm quatro patas.

31) *The dogs* have four legs.*

Os cachorros têm quatro patas.

Além das possibilidades mencionadas, há também no Português Brasileiro (PB) a possibilidade do SN nu singular, como no exemplo a seguir (SCHMITT; MUNN, 1999, p. 7), que, apesar de estar gramaticalmente no singular, possui leitura de plural:

32) *Cachorro* gosta de gente.

Contrastando com os nomes genéricos há os *nomes referenciais*, [+ específicos], que ocasionalmente também são [+ definidos]. No PB, esses nomes são normalmente acompanhados de um determinante definido (33) ou demonstrativo (34). Apesar de os nomes referenciais serem necessariamente [+ específicos], os nomes genéricos não devem ser confundidos com os [- específicos] (LYONS, 1999, p. 186). Há uma diferença entre eles, apresentada a partir das duas leituras do exemplo (35)⁸:

33) *O rapaz* da loja me pareceu muito simpático.

34) Comprei *esse livro* ontem.

35) *An Indian* smokes a pipe every night.

Um indiano fuma um cachimbo toda noite.⁹

⁷ Tradução nossa dos exemplos (28), (29), (30) e (31).

⁸ Lyons (1999, p. 186).

⁹ Tradução nossa.

Na leitura [- específica] da frase (35), fumar cachimbo todas as noites pode ser uma ação realizada por indianos diferentes. Na leitura genérica, o ato de fumar cachimbo é percebido como uma característica que faz parte do que significa ser indiano. Os traços [- específico] e [+ genérico] não são, portanto, equivalentes. Há, por fim, a leitura [+ específica], na qual um indiano particular é um fumante noturno. Como mencionado anteriormente, apenas os SNs com leitura [+ específica] [+ definida] foram incluídos no presente estudo.

A noção de **contabilidade e número**, discutida em seguida, foi inserida pois se mostrou relevante em estudos sobre os SNs nus em variedades do português¹⁰.

2.1.4 Contabilidade e número

Embora não haja uma seção específica que relacione a definitude à contabilidade e ao número do SN em Lyons (1999), serão apresentadas considerações pertinentes ao tema, identificadas no decorrer da leitura, além de reflexões de outros autores sobre essa relação.

De acordo com o Houaiss (2009), **contável** é o que é capaz de ser contado, de entrar na oposição singular/plural, como ocorre com coisas materiais atomizáveis e descontínuas (*apartamentos, pessoas*) e coisas abstratas que podem ser individualizadas (*poderes, governos, emoções*). **Não-contáveis**, ou nomes *de massa*, são aqueles que, diferentemente dos primeiros, não podem ser atomizados, transformados em unidades, como *areia e água*. As noções de contabilidade e número estão intrinsecamente ligadas, uma vez que os nomes de massa só se apresentam no singular, enquanto os contáveis podem aparecer também no plural. A pluralização é, portanto, *um reflexo gramatical da contabilidade* (CAMACHO; PEZATTI, 1996, p. 61)¹¹; o traço [- plural], por sua vez, pode estar presente em usos contáveis e não-contáveis. Os exemplos a seguir, trazidos por Camacho e Pezatti (1996, p. 62) ilustram essa afirmação:

36) Vi *um frango* no quintal. [+ contável] [- plural]

37) Vi *três frangos* no quintal. [+ contável] [+ plural]

38) Vai ter *frango* no jantar. [- contável] [- plural]

¹⁰ PABH (BAXTER; LOPES, 2009; NEVES, 2019) e PA (FIGUEIREDO, 2019).

¹¹ Grifo nosso.

O autor e a autora afirmam que uma forma de identificar se um SN é [\pm contável] é a partir da sua replicabilidade. Um nome contável é caracterizado por ser replicável, i.e., há um ponto em que sua instância se esgota e, para incrementá-la, se acrescenta outra (*um frango, dois frangos...*). Os nomes de massa não possuem esse limite e podem ser, por isso, expansíveis. Pode-se pensar que, no jantar (38), alguém pode gostar muito do prato servido e pensar que queria *mais frango*, mas tem *pouco frango* na mesa.

Ao discutir esse tema na língua inglesa, Lyons (1999, p. 10) afirma que os artigos definidos podem ocorrer igualmente nos contextos de nomes contáveis – no singular ou no plural – e nos nomes de massa. Os exemplos a seguir (p. 6 e 10) representam os três contextos, respectivamente:

39) Pass me *the hammer*, will you?

Me passa *o martelo*, por favor?

40) We've just been to see John race. The Queen gave out *the prizes*.

Acabamos de ver John correndo. A rainha distribuiu todos *os prêmios*.

41) We went to the local pub this lunch time. They've started chilling *the beer*.

Nós fomos para o pub local na hora do almoço. Eles já tinham começado a gelar *a cerveja*.

No PB também há essas três possibilidades, como pode-se ver a partir das traduções feitas. Há, porém, mais uma relação entre número e contabilidade do SN no PB, através do SN nu. Como ocorre no turco, a ausência de marca de plural nos substantivos contáveis não implica sua singularidade (LYONS, 1999, p. 185). Os nomes são, portanto, neutros quanto ao número, e a ausência da marca de plural pode representar uma classe:

42) Anna ama *bolo*.

No caso do exemplo (42), o SN é singular com leitura de plural. O *bolo* representa, assim, toda a classe, sugerindo que Anna gosta de qualquer 'bolo'. Apesar de não termos incluído os SDs genéricos no *corpus* deste trabalho, o uso de SDs sem marcação de plural com leitura de plural [+ específica][+ definida] é encontrado no português. Diante disso, incluiu-se a **contabilidade e**

número como uma das variáveis relacionadas ao número no SD. Mais uma noção relacionada à definitude, de acordo com Lyons (1999), é a **animacidade**, que será apresentada a seguir.

2.1.5 Animacidade

Os SNs caracterizados como [+ animados] são os que se referem a seres vivos, *pessoas e animais*, como *a mulher*, na frase (43), e se opõem aos [- animados], como *o café*, em (44). Ao tratar da relação entre animacidade e definitude, Lyons (1999, p. 215) apresenta uma hierarquia de animacidade, sobre a qual comenta que “os referentes humanos são, em geral, mais salientes na percepção humana do que os não-humanos”¹², favorecendo, portanto, a realização do SD nu.

43) *A mulher* leva as compras no carrinho.

44) *O café* que ele trouxe estava frio.

Com base nessa ideia, Baptista (2007) analisa as implicações da saliência do [+ humana] no Crioulo Caboverdiano (CCV) e constata que há uma relação entre definitude, animacidade e flexão de plural nessa língua (BAPTISTA, 2007, p. 86). No Santome (ST), Alexandre e Hagemeyer (2007, p. 50) apontam que o substantivo [+ humano] geralmente tem leitura singular quando realizado como SN nu. Ademais, enquanto os nomes [- humanos] definidos e plurais geralmente aparecem em estruturas compostas por *inen*¹³ + *nome* + *constituente marcador de referencialidade*, os nomes [+ humanos] são os únicos que ocorrem acompanhados exclusivamente do *inen* (ALEXANDRE; HAGEMEIJER, p. 41). Baxter e Lopes (2009) e Neves (2019) também investigam o efeito da animacidade no PABH de forma mais ampla, através dos traços [+ animados]. Nesta pesquisa, seguindo os dois últimos autores, os SNs são classificados como [+ animados] e [- animados], a fim de observar o efeito da animacidade, de forma ampla, sobre a realização dos nomes nus.

Lyons (1999) indica, ainda, que, ao tratar dos diferentes **tipos de posse** e associá-los à noção de definitude, o contraste entre as posses alienável e inalienável se mostra relevante para esse debate e será exposta a seguir.

¹² Tradução nossa.

¹³ *Inen* é um marcador que atribui simultaneamente traços [+ definido][+ plural] aos nomes que acompanha (ALEXANDRE; HAGEMEIJER, 2007, p. 39-40).

2.1.4 Tipos de posse

Lyons (1999, p. 128) inicia a discussão sobre posse afirmando que as línguas distinguem de diversas formas os tipos de relação entre possuidor e coisa possuída. Dentre elas, a que parece ter influência na definitude é a que diferencia a posse **alienável** da **inalienável**, sendo esta última representada referentes ligados ao possuidor de forma mais intrínseca, como *relações familiares* e *partes do corpo*. Em várias línguas em que os SNs recebem um morfema indicando essa diferença, na marcação da posse inalienável “a estrutura geralmente é morfologicamente mais simples ou o possessivo está de alguma forma estruturalmente mais próximo do substantivo principal”¹⁴ (LYONS, 1999, p. 128). Esse é o caso do swahili (língua bantu mais falada), que tem um processo opcional de redução quando a posse é inalienável, integrando o possessivo ao núcleo: *mwenzi wako* (cônjuge seu) “seu cônjuge” pode ser reduzido para *mwenzio*. Um outro exemplo (p. 129) ocorre no dyirbal (língua aborígine da Austrália): apesar de existirem dois morfemas genitivos para posse presente e passada, ambas alienáveis, seu uso é dispensado quando se trata de uma posse inalienável. Essa distinção não ocorre no PB, como pode-se constatar nos exemplos (45) e (46), respectivamente alienável e inalienável:

45) *Meu carro* quebrou no meio do caminho.

46) Encontrei *minha mãe* na padaria.

Sobre os tipos de posse discutidos, Neves (2019, p. 39) afirma que, se o que é inalienável está intrinsecamente relacionado a quem o possui, e o ouvinte conhece o possuidor, a familiaridade com o que é possuído é, portanto, maior do que com qualquer coisa alienável, i.e., que não seja tão intimamente relacionada ao possuidor. Assim, parece haver relação entre os traços [+ definido] e [- alienável], o que pode favorecer a ocorrência de SNs nus.

Elucidamos, até aqui, alguns conceitos semânticos relacionados à definitude, apresentados sobretudo por Lyons (1999), que são fundamentais para a formulação das nossas hipóteses de trabalho na análise quantitativa dos dados. A seguir, apresentaremos, aspectos semânticos e sintáticos basilares para o entendimento da estrutura dos SD nus.

¹⁴ Tradução nossa.

2.2 UMA PERSPECTIVA SEMÂNTICO-SINTÁTICA DOS SDs NO PORTUGUÊS

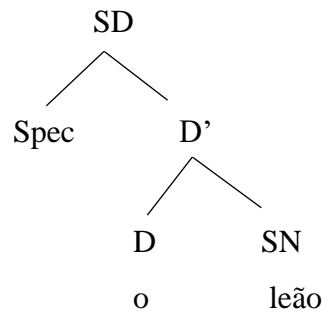
A partir da Hipótese do SD, apresentada por Abney (1987), será apresentada a relação entre as estruturas do SN e do SD nas línguas em geral. Sob essa perspectiva, o D funciona como uma projeção funcional que domina o SN, concedendo-lhe, assim, o *status* de argumento. Chierchia (1998), em seguida, por meio do Parâmetro do Mapeamento Nominal, classifica as línguas a partir da presença ou ausência de dois traços semânticos, $[\pm \text{arg}]$ e $[\pm \text{pred}]$. Além de afirmar que não apenas os SDs, mas também os SNs, podem ser argumentos nas línguas em geral. Chierchia (1998) observa que o que as diferencia é *se e sob quais circunstâncias* isso acontece. Por fim, Schmitt e Munn (1999) demonstram que o PB não se enquadra no Parâmetro de Chierchia (1998), e apresentam a proposta de classificação dos SNs nus singulares dessa língua como SDs com determinantes vazios e sem número.

2.2.1 A Teoria X-Barra e a Hipótese do SD

Para o estudo da sintaxe do uso dos determinantes, este trabalho baseia-se no tema também à luz da Gramática Gerativa, sobretudo da Teoria X-Barra (CHOMSKY, 1970)¹⁵. Essa teoria propõe que há níveis lexicais (V, N, P, A, Adv) e sintagmáticos (VP, NP, PP, AP e AdvP) e, entre eles, há projeções intermediárias (V', N', P', A', Adv'). Desde sua origem, na década de 70, a Teoria X-Barra passou por várias reformulações, dentre elas a feita por Abney (1987), que estabeleceu a projeção das categorias funcionais. Em sua análise, Abney (1987) propõe que o D possui as propriedades de um elemento funcional (ABNEY, 1987, p. 49) e, através da comparação da estrutura do SN com a das sentenças¹⁶, afirma que o SN é dominado pela categoria funcional SD e o seu núcleo (D) atribui referencialidade ao SD, dando-lhe o *status* de argumento. Os argumentos são, portanto, denominados SD, cuja estrutura é assim representada:

¹⁵ Chomsky (1957, p. 13) afirma que o principal objetivo da análise linguística de uma língua é separar as suas sequências gramaticais das não-gramaticais e estudar a estrutura das que são gramaticais. A Teoria X-barra (Chomsky, 1970) traz uma estrutura representativa dos constituintes das línguas, evidenciando a natureza deles, as relações internamente estabelecidas e a forma como se hierarquizam para formar sentenças.

¹⁶ Segundo Abney (1987), assim como a projeção InflP domina VP, o DP domina o NP.



N teoria X-Barra, o SD é a projeção funcional do núcleo do sintagma nominal. O SD é formado pela concatenação de um núcleo D, com, no mínimo, um nome. Os núcleos do sintagma SD são determinantes que, por hipótese, c-selecionam como complemento lexical um SN.

2.2.2 O Parâmetro do Mapeamento Nominal de Chierchia (1988)

Chierchia (1988) estabelece, através de uma interface semântico-sintática, o Parâmetro do Mapeamento Nominal (PMN), uma classificação tipológica das línguas naturais que as distribui de acordo com o que suas projeções nominais máximas podem denotar. Ao analisar, entre outros aspectos, a variação no uso do determinante, este Parâmetro tem se mostrado relevante para os estudos de SNs nus em diversas línguas.

Opondo-se à visão tradicional de que os Ns e SNs são sempre predicados (ou seja, denotam propriedades¹⁷) e que apenas Sintagmas Determinantes (SDs) podem ser argumentos (ou seja, denotam espécies¹⁸), Chierchia (1988, p. 343-344) parte do princípio de que os SNs nus também podem ser argumentos, e a questão passa a ser *se e sob quais condições* isso acontece.

A denotação de substantivos pode variar entre os idiomas, e essa variação pode ser responsável pelas diferentes distribuições de argumentos nominais simples (CHIERCHIA, 1998, p. 344). A partir do PMN, busca-se sistematizar essa variação e se propõe que há três tipos de línguas naturais, de acordo com as combinações dos traços semânticos [\pm arg] e [\pm pred]. O PMN pode ser assim sintetizado (SCHMITT; MUNN, 1999, p. 341):

¹⁷ Müller (2002, p. 326) explica que *propriedades* se referem a funções que caracterizam um conjunto de indivíduos em um mundo, como, por exemplo, *professor*.

¹⁸ Espécies, de acordo com Carlson (1977 apud MÜLLER, 2002, p. 326), são as “entidades do mundo”, como *Fernando Henrique*.

47) O Parâmetro do Mapeamento Nominal:

a. [+arg, -pred] (exemplo: chinês)

- argumentos nus generalizados;
- todos os nomes são de massa;
- sem morfologia de plural;
- sistema de classificadores generalizado.

b. [-arg, +pred] (exemplo: francês)

- sem nomes nus na posição de argumento;
- distinção entre nomes contáveis e de massa;
- morfologia de plural.

c. [+arg, +pred] (exemplo: inglês)

- nomes nus de massa e plurais nus na posição de argumento;
- sem nomes nus singulares contáveis;
- morfologia de plural.

[-arg,-pred] (inexistentes).¹⁹

O PMN, cujos exemplos se baseiam sobretudo no inglês (representando as línguas germânicas) e italiano (representando as línguas românicas), não abrange todas línguas naturais, como, por exemplo, o PB. A impossibilidade do enquadramento do PB no Parâmetro de Chierchia (1998), tendo em vista o comportamento dos seus SNs nus, é apontada por Schmitt e Munn (1999).

2.2.3 A insuficiência do Parâmetro do Mapeamento Nominal para classificar o PB e a proposta de Schmitt e Munn (1999)

Ao revisarem os aspectos principais do Parâmetro Nominal de Chierchia (1998), Schmitt e Munn (1999) constatam que o PB não se enquadra em nenhum dos tipos de língua incluídas na proposta de Chierchia. Preliminarmente, os autores mostram que o PB se assemelha ao inglês

¹⁹ Tradução nossa.

quanto aos nomes nus plurais, exemplo (48), e nomes de massa, exemplo (49), que podem ocorrer na posição de argumento – o que não é permitido na maioria das línguas românticas (SCHMITT; MUNN, 1999, p. 339). Diferente do inglês, porém, o PB também permite nomes contáveis singulares como argumentos, exemplo (50), ao contrário do sugerido por Chierchia (1998)²⁰:

48) Crianças leem *revistinhas*.

Children read *comic books*.

49) *Ouro* é caro.

Gold is expensive.

50) Ele comprou *computador*.

He bought *a computer/computers*.

Posteriormente, Schmitt e Munn (1999, p. 349) investigam mais detalhadamente por que não é possível enquadrar o PB em nenhum dos tipos de língua propostos por Chierchia (1998)²¹. Segundo os autores, o Parâmetro estabelece que as línguas [+arg, -pred], como o chinês, não distinguem morfologicamente singular/plural, nem diferenciam os nomes contáveis e os nomes de massa. O PB, por outro lado, apresenta morfologia distintiva de plural e diferencia os nomes contáveis dos de massa (SCHMITT; MUNN, p. 349). Chierchia (1998)²² assume, ainda, que (i) as línguas que possuem as distinções mencionadas somente permitem que plurais e nomes de massa denotem tipo, e que (ii) nas línguas [+arg, -pred] apenas os nomes de massa são [+arg] – o que também não se aplica ao PB. Por fim, Schmitt e Munn (1999, p. 350) elucidam que o PB admite nomes singulares nus como argumentos de verbo, o que não é possível nas línguas classificadas como [-arg, +pred] e [+arg, + pred] no Parâmetro Nominal.

Após constatarem que o sistema de Chierchia (1998) é insuficiente para explicar o uso dos nomes nus singulares no PB, os autores sugerem que o mais adequado é considerá-los como SDs com determinantes nus (SCHMITT; MUNN, 1999, p. 351). Para isso, apresentam evidências de que, embora sejam sintaticamente singulares, esses nomes não têm especificação semântica de

²⁰ “Tanto no alemão como nas línguas românticas, argumentos singulares nus são totalmente impossíveis (se o nome não for de massa)” (CHIERCHIA, 1998, p. 341, tradução nossa).

²¹ apud Schmitt e Munn (1999, p. 349).

²² apud Schmitt e Munn (1999, p. 349).

número: a depender do contexto, os SNs nus singulares se comportam ora como semanticamente singulares, ora como semanticamente plurais (SCHMITT; MUNN, 1999, p. 351).

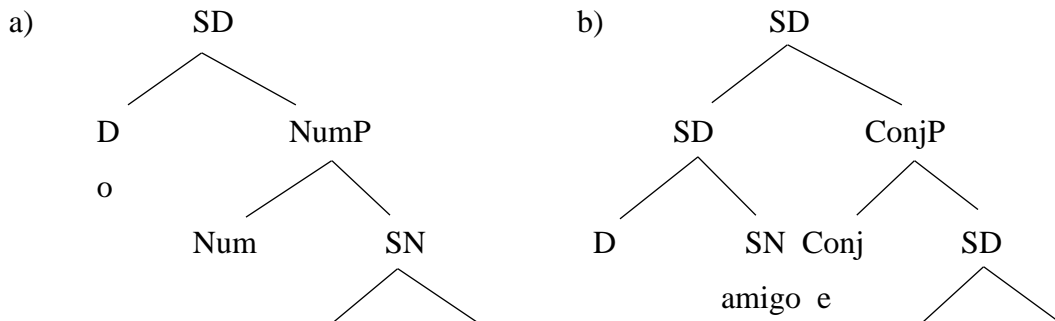
Se os nomes nus singulares não têm especificação semântica de número e também não têm leitura de massa, pode-se pensar que eles são SNs, não SDs, mas Schmitt e Munn (1999, p. 352) não acreditam que esse seja o caso. Segundo eles, se os SNs nus singulares denotam predicado, a junção de dois SNs nus produz outro predicado, como é o caso da frase do exemplo (51), em que se juntam dois NPs sob apenas um determinante. Nesse caso, a leitura admitida no PB é a de que se trata de uma pessoa que é, ao mesmo tempo, amigo e parente. Esse pressuposto não dá conta, porém, de ocorrências como o exemplo (52), em que só se pode interpretar que os SNs se referem a ‘pessoas que são amigos’ e ‘pessoas que são parentes’, separadamente. Afinal, se os SNs nus singulares denotam somente predicados, a junção de dois SNs deveria permitir somente a formação de um predicado conjunto, o que não ocorre nesse exemplo. No exemplo (53), por fim, as duas interpretações são possíveis: (i) as mesmas pessoas que são amigos, também são parentes, ou (ii) ele encontrou pessoas que são seus amigos e pessoas que são seus parentes (SCHMITT; MUNN, 1999, p. 352).

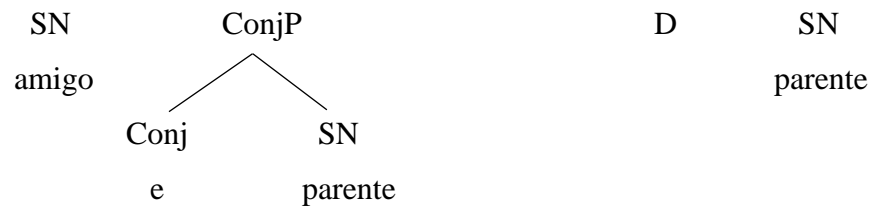
51) Ele encontrou *o amigo e parente* no aeroporto.

52) Eu encontrei *amigo e parente* no aeroporto.

53) Ele encontrou *os amigos e parentes* no aeroporto.

Pode-se dar sentido a essa distinção ao se assumir que há um DP dividido que contém NumP, onde se localiza o número semântico, segundo Schmitt e Munn (1999, p. 352). Assim, as duas leituras decorrem das diferentes estruturas: a leitura do predicado conjunto, como em (51), se dá a partir da combinação do SN, apresentado em (a), ao passo que a leitura da existência de dois indivíduos diferentes, exemplificada em (52), surge a partir da combinação NumP ou SD, mostrada em (b):





Assumindo que é esse o caso, os autores propõem que a combinação de nomes nus não pode ser uma simples combinação de SNs, e que nomes singulares nus são SDs com determinantes vazios e sem número (SCHMITT; MUNN, 1999, p. 353).

Depois de apresentarem evidências de que o PB não se enquadra no Parâmetro do Mapeamento Nominal de Chierchia (1998), sobretudo diante da insuficiência para explicar os SNs singulares nus nessa língua, Schmitt e Munn (1999) apresentam uma estrutura alternativa. Através da interpretação dos SNs singulares nus como SDs, alguns dos seus contextos de uso podem ser explicados. Os autores usam, porém, um número reduzido de dados, cumprindo o intuito de ilustrar a argumentação teórica. A seguir, serão apresentados estudos quantitativos e descritivos dos SNs nus em variedades do português: o PB (WALL, 2013), o PABH (BAXTER; LOPES, 2009; NEVES, 2019) e o PA (FIGUEIREDO, 2019).

2.3. PESQUISAS SOBRE SDs NUS DE REFERÊNCIA DEFINIDA EM VARIEDADES DO PORTUGUÊS

A formação de um sistema de referências que inclui a ausência do artigo definido pode estar, em algum nível, ligada a questões de economia de sistema orientadas pelo grau de previsão de referência atribuído pelos contextos sintático-discursivos e pragmáticos (BAXTER; LOPES, 2009). Nesta seção, são apresentados estudos em variedades do português que analisam quantitativamente e descrevem os SNs nus nos diversos contextos em que são realizados: variedades do português urbanas (WALL, 2013) e rurais – o Português Afro-Brasileiro de Helvécia (BAXTER; LOPES, 2009), e o Português de Almojarife (FIGUEIREDO, 2019).

2.3.1 Os SDs nus [+ específicos][+ definidos] no Português Brasileiro (WALL, 2013)

Ao constatar que a maior parte dos estudos sobre SNs nus no português se restringe a analisá-los em seu uso genérico e indefinido, Wall (2013, p. 223-225) decide investigar os SNs nus com leitura definida e específica no PB. Ele afirma que, ao contrário de outras línguas românicas que têm restrições bastante limitantes quanto ao uso dos SNs singulares nus²³, no PB, a situação é diferente²⁴ – embora não ocorram em qualquer contexto, aleatoriamente (WALL, 2013, p. 226). A fim de analisar a sistematicidade do uso dos SNs singulares nus, ilustrados pelos exemplos (54) e (55)²⁵, o autor analisa dois *corpora*: (i) um *corpus* aberto, com dados da língua escrita e falada do NURC; (ii) um *corpus* fechado, em que 37 participantes narram histórias enquanto acompanham um livro de figuras. (WALL, 2013, p. 233).

54) Eu não gosto do quintal [...] *quintal* é cheio de terra.

55) Ela tem uma filha. [...] *Filha dela* tem sete anos.

Na análise do *corpus* aberto, tanto na fala (75 dados) quanto na escrita (24 dados) os SNs singulares nus de referência específicas e definidas ocorrem predominantemente²⁶ na posição de sujeito pré-verbal, como nos exemplos mostrados. As ocorrências de SNs singulares nus na posição pós-verbal não são suficientes para fazer generalizações, apenas a constatação de que elas são possíveis em certos contextos (WALL, 2013, p. 237). No que diz respeito à modalidade oral, Wall (2013, p. 238) observa dois fatos importantes: (i) todos os SNs singulares nus realizados possuem referência específica e definida, e (ii) a maioria dos referentes desses SNs já tinham sido introduzidos na conversa diretamente (46/75), por anáfora, ou indiretamente (72/75), por referência cruzada. Houve, ainda, alguns referentes que estavam fisicamente presentes no momento da enunciação²⁷ (WALL, 2013, p. 238). A modalidade escrita, no *corpus* aberto, além de estar sujeita a pressões normativas diferentes daquelas que regem a fala, mostra uma distribuição de dados assistemática e pouco confiável, devido à pequena quantidade de dados (WALL, 2013, p. 238-239). No *corpus* fechado, por sua vez, embora a frequência de de SNs singulares nus com leitura específica e

²³ Os SNs singulares nus, nessas línguas, não podem exercer a função de argumento (WALL, 2013, p. 226).

²⁴ “No português do Brasil, diferentemente do que ocorre nas outras línguas românicas e no português europeu, essa construção é frequente na fala cotidiana” (BRAGA ET AL., 2008, p.110 apud WALL, 2013, p. 228).

²⁵ Exemplos mencionados por Wall (2013, p. 236).

²⁶ Na fala, foram 69 sujeitos pré-verbais, 1 objeto pré-verbal e 5 objetos pós-verbais. Das 24 ocorrências escritas, 11 foram SNs nus definidos: 7 sujeitos pré-verbais, 1 sujeito pós-verbal, 2 objetos diretos pós-verbais e 1 objeto indireto pós-verbal (WALL, 2013, p. 235).

²⁷ O autor não especifica quantos referentes estavam fisicamente presentes na enunciação.

definida não corresponda a uma grande amostra (158/6333), 156 ocorrências correspondem a sujeitos pré-verbais e apenas 2 a objetos pós-verbais (WALL, 2013, p. 240-241). Observa-se, a partir dos resultados, que os dados de SNs singulares específicos e definidos não se enquadram nas abordagens teóricas convencionais para os SNs singulares nus no PB, nas quais se assume que o artigo definido é o elemento responsável pela correspondência entre os tipos semânticos (a realização do operador *iota*)²⁸ e as projeções sintáticas (o núcleo do DP) (WALL, 2013, p. 242).

Por fim, Wall (2013, p. 250-251) salienta que, devido a um ‘ponto cego’ teórico-metodológico nos estudos anteriores, eles falharam em identificar um tipo de SNs singulares nus no PB: os que têm leitura específica. Embora esses SNs se comportem mais como exceção do que como regra, eles não devem ser excluídos dos conjuntos de dados. Apenas uma combinação de diferentes abordagens teóricas é capaz de oferecer uma visão mais completa desse fenômeno (WALL, 2013, p. 250).

2.3.2 Os SDs nus no Português Afro-Brasileiro de Helvécia

Os dois estudos do PABH inseridos nesta seção têm particularidades específicas. Baxter e Lopes (2009) estudam a variação no uso do artigo definido através de uma variável dependente binária (presença e ausência do artigo definido). Neves (2019), por sua vez, na análise da variação dos determinantes nos objetos diretos [+ específicos][+ definidos], utiliza uma variável trinomial, incluindo os demonstrativos não-dêiticos.

2.3.2.1 A variação na realização do artigo definido nos SNs [+ definidos] de Helvécia (BAXTER; LOPES, 2009)

Baxter e Lopes (2009) investigam em dados do Português Afro-Brasileiro de Helvécia (PABH) a variação no uso dos determinantes dos SNs de referência definida, fenômeno com importante função nos crioulos de base lexical portuguesa²⁹. Para isso, analisam um *corpus* de onze entrevistas sociolinguísticas, distribuídas entre falantes de 4 faixas etárias: faixas 2 (35 a 45 anos), 3 (60 a 70 anos), 4 (mais de 80 anos), e ‘faixa 5’ (um informante de mais de 100 anos). 2.655 SNs

²⁸ Dobrovie-Sorin e Pires de Oliveira (2008, p. 112) apud Wall (2013, p. 242).

²⁹ Lucchesi (1993), Alexandre e Hagemeyer (2007) e Baptista (2002, 2007).

de referência definida foram realizados, dos quais 80,2% são com o artigo definido (BAXTER; LOPES, 2009, p. 320).

As variáveis selecionadas como condicionantes da realização do artigo definido foram, em ordem de seleção: *faixa etária*, *referência do SN*, *número do SN*, *presença de outro constituinte capaz de induzir referência definida*, *interação com o determinante do SN correferencial prévio*, *função sintática do SN*. As únicas variáveis consideradas estatisticamente não significativa pelo GOLDVARB foram a *animacidade* e *contabilidade do substantivo* (BAXTER; LOPES, 2009, p. 322).

Na análise do efeito da *faixa etária* sobre a realização do artigo definido, é constatado um nítido perfil aquisicional: as faixas 4 e 5, em fase inicial de aquisição, realizam SNs nus de forma significativa, enquanto nas faixas 2 e 3 observa-se um uso gradual das formas mais próximas do padrão (BAXTER; LOPES, 2009, p. 322-323). Ao avaliarem o efeito da *referência do SN*, confirma-se a hipótese de que, nos SNs que possuem referentes claramente identificáveis, como no exemplo (56)³⁰, o conhecimento compartilhado pelos interlocutores favorece o uso do artigo definido. Seu uso é desfavorecido, por outro lado, quando a referência é [+ única] (BAXTER; LOPES, 2009, p. 323).

56) Tava oiano na lua assim ói, *lua* tá traz do sol uma bandinha.

57) Num contano *as mulhê* não.

Quanto ao *número do SN*, o SN singular mostra-se ligeiramente desfavorável à realização do artigo definido, ao passo que o SN plural, exemplificado em (57), o favorece fortemente. Ao cruzar esse grupo com a *faixa etária*, observa-se que o SN plural é favorável ao artigo definido em todas as faixas, enquanto o singular só o favorece na faixa 2 (BAXTER; LOPES, 2009, p. 323-324).

O efeito da *presença de outro constituinte capaz de induzir referência definida*, cujas variantes estão exemplificadas em (58-64), mostra, por um lado, que o uso do artigo definido é inibido diante do possessivo³¹ e do sintagma preposicional. Por outro lado, essa inibição não ocorre

³⁰ Os exemplos mostrados aqui são os utilizados pelos autores.

³¹ Os autores salientam que “os possessivos pré-nominais têm um comportamento muito diferente daquele dos outros modificadores contemplados aqui e são capazes de atribuir um grau de referência definida suficiente para dispensar a presença do artigo definido, uma característica geral no português brasileiro.” (BAXTER; LOPES, 2009, p. 326)

na presença de outros tipos de constituintes, como o adjetivo pré-nominal, o adjetivo pós-nominal e a oração relativa (BAXTER; LOPES, 2009, p. 326).

58) Nenhum outro modificador: *ficô eu pra contá o caso.*

59) Oração relativa: *a casa que eu morava é lá.*

60) Adjetivo pré-nominal: *é a merma coisa.*

61) Adjetivo pós-nominal: *a terra vermelha, ela aguenta.*

62) Sintagma preposicional: *eu vi distância do caminho.*

63) Possessivo: *o meu pai ajuntô cum ela.*

64) Advérbio locativo: *a bica ali secô.*

Uma perspectiva aquisicional é capaz de explicar a maior redução do uso do artigo definido na presença dos advérbios locativos, em relação aos outros constituintes, visto que no PABH os locativos cumprem o papel de demonstrativo, atuando como determinante (BAXTER, LUCCHESI E GUIMARÃES, 1997, p. 13-14)³². Ademais, a coocorrência do artigo definido plural e de modificadores pós-nominais mostra semelhanças com o ST (ALEXANDRE, HAGEMEIJER, 2007), em que são encontrados substantivos modificados simultaneamente pelo marcador [+plural][+definido] *inen* e por outros modificadores pós-nominais, como a oração relativa, o adjetivo e o possessivo (BAXTER; LOPES, 2009, p. 327).

Os resultados da *interação com o determinante do SN correferencial prévio* mostram uma interação complexa entre os determinantes do SN correferente prévio (SN1) e do SN correferente subsequente (SN2): quando há o artigo definido ou o determinante \emptyset no SN1, a repetição desse determinante é favorecida no SN2, como nos exemplos (65) e (66), respectivamente. A realização do demonstrativo no SN1, no entanto, favorece a ausência do artigo definido no SN2, como exemplificado em (67).

65) Artigo definido: quando eu tava com *o machado*, quando eu bati *o machado* no pau, quando bateu fez <taco>!

66) Sem artigo definido: Essas menina é... é pa barrê *terreiro*. [...] poque *terrêro* era grande.

67) Demonstrativo: Qu'ocê num contá *esse garucho*. Pá *garucho* levá!

³² apud Baxter e Lopes (2019, p. 326).

Os autores afirmam que, devido à atribuição dos valores referenciais de dêixis e definitude pelo demonstrativo, “a referência do SN está suficientemente fundamentada para dispensar o determinante na ocorrência subsequente desse SN” (BAXTER; LOPES, 2009, p. 327-328).

O efeito do último grupo analisado, *a função sintática do SN*, revela que, enquanto o artigo definido só é fortemente favorecido na função de adjunto adnominal, em várias funções sintáticas ele é desfavorecido. Dentre todos os fatores que inibem o uso do artigo definido, o sujeito e o objeto direto são os que o fazem de forma mais leve e estão ilustrados respectivamente em (68) e (69). O SN independente, por sua vez, mostra um peso relativo neutro. (BAXTER; LOPES, 2009, p. 328-329).

68) *Mãezinha de nós* caregava nós assim.

69) Eu vi *distância* do caminho.

O perfil geral dos resultados revela, por um lado, que o PABH mostra um perfil de aquisição do artigo definido pleno, ao passo que apresenta evidências de uma fase anterior no dialeto, quando o SN nu definido teria sido mais frequente; por outro lado, são apresentadas características específicas desse dialeto, além de algumas semelhanças com a tipologia do SN sem artigo definido no CCV e no ST. (BAXTER; LOPES, 2009, p. 330)

2.3.2.2 O uso dos SDs nus objetos no PABH: um olhar direcionado para os dados referenciais (NEVES, 2019)

No intuito de avaliar a variação no uso dos SDs na função de objetos diretos [+ específicos][+ definidos] no PABH, Neves (2019) analisa dados de 18 entrevistas sociolinguísticas do projeto intitulado *Creole-like traces in Afro-Brazilian rural communities*, dirigido pelo Prof. Dr. Alan Baxter e patrocinado pela *La Trobe University Faculty of Humanities Research Committee* e pelo *Australian Research Council*. As entrevistas estão distribuídas igualmente entre os sexos (masculino/feminino) e faixas etárias: faixas 1 (20 a 40 anos), 2 (41 a 60 anos) e 3 (> 65 anos) (NEVES, 2019, p. 155).

Foram quatro os grupos selecionados como estatisticamente relevantes para o condicionamento da variável em estudo, em ordem de seleção: *faixa etária, presença de outro constituinte que marque referencialidade, familiaridade e número no SD*. Os resultados do efeito da *faixa etária* apontam para um sistema de determinante em desenvolvimento, visto que há uma relação diretamente proporcional entre as faixas etárias e o favorecimento do uso do determinante Ø (NEVES, 2019, p. 171-172). Na análise ternária (determinante Ø, demonstrativo e artigo definido), observa-se um panorama mais detalhado dessa variação: os falantes mais idosos se aproximam mais da gramática em situação de contato, com preferência pelo Det Ø (> demonstrativo > artigo definido); na faixa 2, a leitura definida é expressa principalmente pelo determinante foneticamente realizado mais saliente, o demonstrativo³³ (> artigo definido > Det Ø); e na faixa 1, há a preferência pelo artigo definido (> demonstrativo > Det Ø). A quebra do isolamento vivido pela comunidade a partir do fim da colônia, em 1888, explica o perfil dos resultados registrados pelo estudo em tempo aparente, visto que o acesso a meios de transporte e de mídia (como televisão e rádio), suscitou o contato dos falantes do PABH com outras variedades do português, inclusive as mais padronizadas. (NEVES, 2019, p. 173-174)

Na análise da *presença de outro constituinte que marque referencialidade*, por um lado, quatro tipos de constituintes (sintagma preposicional, locativo, possessivo e adjetivo, ilustrados em (70-73)) favorecem o determinante Ø, confirmando o caráter econômico do PABH e a hipótese de que, na presença de outro constituinte no SD, o determinante foneticamente realizado se torna desnecessariamente redundante. A presença das relativas e dos apostos (exemplificados em (74) e (75)), por sua vez, desfavorece o SD nu e parece ser ‘a porta de entrada’ dos determinantes foneticamente realizados (NEVES, 2019, p. 176).

70) Sintagma preposicional: ...qu'eu tava sentino, assim, *negóço duma fraqueza...*

71) Locativo: ...fez *operação* lá no...no... *em Belo Horizonte*.

72) Possessivo: Eu peguei *carta dela*.

73) Adjetivo: Anjo de Guarda dele é fôte, topô *onça morto*.

74) Relativas: Eu mostro *o serviço que ele já fez*.

³³ “Essa é uma solução típica em processos de aquisição em situação de contato, pois o artigo definido é foneticamente fraco e apresenta dificuldade para ser identificado e decodificado em termos de função. No caso do português, isso acontece porque a sua realização fonética é coincidente com a de uma preposição da língua e, muitas vezes coincide com a vogal final ou inicial de uma palavra” (NEVES, 2019, p. 174).

75) Apostos: Aí ele mandô *a mãe, essa véia aí mesmo*.

Quanto à *familiaridade*, característica dos SDs [+ específicos][+ definidos], a análise mostra que os fatores *catáfora* e *referência cruzada* favorecem o determinante \emptyset , enquanto os fatores *anáfora* e *situacional* o desfavorecem (NEVES, 2019, p. 181).

76) Anáfora: Que ela tá bateno pá botá *esse papel*. (o papel já havia sido citado anteriormente)

77) Referência Cruzada: quando ele tirô *o ponto*. (os pontos da operação)

78) Situacional: ...eu arrumei *essa menina* aqui, ó.

79) Catáfora: ...quando pensô que não, começô *poblema de pressão*...

Ao tratar da *noção de número no SD*, Neves (2019, p. 184) observa preliminarmente que possibilidade de um nome nu [+ contável] ser tratado como uma entidade individual no PABH evidencia limitações no Parâmetro do Mapeamento Nominal de Chierchia (1998) e mostra semelhanças entre essa língua e o CCV (BAPTISTA, 2007, p. 73-74). Na análise dos dados do PABH, os resultados do efeito desse grupo revelam que o SD *singular*, como no exemplo (80), favorece levemente o determinante \emptyset , ao passo que o desfavorecimento registrado pelo fator *plural* + *leitura de plural*, exemplificados em (81) e (82), parece ser um reflexo da possibilidade da marcação de plural no determinante (NEVES, 2019, p. 184-185).

80) Singular: o médico descobriu *ulça dele*, o ulça dele era já perfurada.

81) Plural: tirô exame, fez, tirô *papéi* da vista.

82) Singular com leitura de plural: Aí tem o *time de Helveça*.

A análise que Neves (2019) apresenta dos grupos considerados estatisticamente não significativos indicam que (i) enquanto a *posse* inalienável favorece o Det \emptyset , a alienável o desfavorece, resultado semelhante ao encontrado no CCV (BAPTISTA, 2007) e condizente com a relação sugerida por Lyons (1999) entre animacidade e os nomes nus; (ii) quanto ao *sexo*, o Det \emptyset é levemente favorecido na fala dos homens, e na fala das mulheres é desfavorecido; (iii) no que concerne à *animacidade*, os nomes [+ animados] desfavorecem o Det \emptyset , ao passo que e os [- animados] o

favorece; (iv) no grupo *contabilidade*, os dois fatores registram valores próximos do ponto neutro. (NEVES, 2019, p. 189-197)

Diante dos resultados apresentados, nota-se que, embora haja resquícios do sistema inicial dos determinantes no PABH (com uma forte presença dos SDs nus), há uma mudança em curso em direção ao uso dos artigos definidos. Os efeitos da *presença de outro material que marque referencialidade* e da *familiaridade* reforçam a importância da interface entre os níveis gramatical, semântico e pragmático na interpretação dos dados. Por fim, a preferência pelo determinante \emptyset nos nomes singulares, dentro do grupo *noção de número no SD*, aponta a peculiaridade do PABH em relação a outras variedades do PB, em que o objeto singular [+Específico] [+Definido] é agramatical. (NEVES, 2019, p. 200-201)

2.3.3 Os SDs nus no Português de Almojarife (FIGUEIREDO, 2019)

Ao investigar a variação no uso dos determinantes definidos no Português de Almojarife (PA), Figueiredo (2019) codificou 13220 SNs, e analisou a variação por meio de uma variável binária (83,8% de SNs nus e 16,2% de inserção do artigo³⁴). Esse estudo não contempla o demonstrativo não-dêitico, mas apresenta três possibilidades de realização do D:

83) Nome nu:

E amigo meu tinha uma mulher.

84) Artigo definido do português ‘o(s)/a(s)’:

É *os meus pai* que me deu essa ideia.

85) Realização do determinante sob a forma do pronome pessoal do português, 3ª pessoa “ele(s)”:

Dentro de dez minuto *ele seôr* já stá morto.

As 18 entrevistas sociolinguísticas que constituem o *corpus* utilizado são igualmente distribuídas por sexo e idade e fazem parte do projeto *Semi-creolization: testing the hypothesis against data from Portuguese-derived languages of São Tomé (Africa)*, dirigido pelo professor Dr. Alan Baxter

³⁴Nesta variante, foram incluídas a realização do artigo definido do português ‘o(s)/a(s)’ e a realização do determinante sob a forma do pronome pessoal do português ‘ele(s)’.

e financiado pela *La Trobe University* (FIGUEIREDO, 2019, p. 361). As variáveis consideradas estatisticamente relevantes pelo GOLDVARB como condicionadoras da inserção do artigo definido foram (i) sociais: sexo, idade e escolaridade; (ii) semânticas: animacidade do substantivo, referência do SN e (iii) estruturais: número do SN e tipo de artigo definido (FIGUEIREDO, 2019, p. 370).

O efeito do *sexo* revela que os homens lideram na inserção do artigo definido, devido, segundo Figueiredo (2019), ao maior acesso à escolaridade e ao contato que eles têm com a comunidade externa; as mulheres, por sua vez, se encontram mais restritas à comunidade e realizam mais o determinante Ø (FIGUEIREDO, 2019, p. 371). Na análise da *faixa etária*, a faixa 1 favorece a realização do artigo fonético, que é levemente desfavorecido pelas faixas 2 e 3. Reportando o nivelamento entre as duas faixas mais avançadas, o autor sugere que há uma mudança consumada ou variação estabilizada na comunidade, “dado o seu isolamento temporal mais prolongado”, que “terá conduzido a uma estabilização da variação” (FIGUEIREDO, 2019, p. 373). A gramática do PA mostra-se, assim, apegada ao sistema da sua fase inicial, com um uso generalizado de SNs nus, uma característica compartilhada com o substrato Santome (FIGUEIREDO, 2019, p. 372-373). Dentre os níveis de *escolaridade*, apenas o mais alto (pós-primário) favorece a realização do artigo definido, grupo em que estão inseridos 50% dos falantes da faixa 1. Há uma associação direta entre idade e escolaridade na comunidade: a faixa 1 é toda escolarizada, e a faixa 3 apresenta um alto nível de analfabetismo. (FIGUEIREDO, 2019, p. 373). Outro fator ao qual Figueiredo (2012, p. 68-69)³⁵ atribui a causa da estabilidade na variação é a composição do corpo docente atuante em São Tomé após a independência: para ocupar os cargos antes ocupados pelos portugueses foram contratados falantes nativos, que têm um padrão mais próximo do português dos alunos do que da variante padrão. Os efeitos das variáveis sociais apontam, assim, para a variação estável ou mudança consumada, confirmando apenas parcialmente as hipóteses formuladas (FIGUEIREDO, 2019, p. 374).

Ao investigar o efeito da *animacidade*, Figueiredo (2019, p. 375) vê que os dados do PA seguem os padrões do Santome (ALEXANDRE E HAGEMEIJER, 2007, p. 49) e do CCV (BAPTISTA, 2007, p. 86): os nomes [+ humanos] inibem a realização do artigo, enquanto os [- humanos] a favorecem. O efeito da *referência do SN* mostra semelhanças entre as variedades de Almojarife e Helvécia e ambos vão ao encontro do que Alexandre e Hagemeyer (2007, p. 53-54)

³⁵ apud Figueiredo (2019, p. 374).

observam no Santome, e Baptista (2007, p. 74), no CCV: a referência [- única] favorece a realização do determinante \emptyset no SN, e a referência [+ única] exerce um peso neutro na inserção (FIGUEIREDO, 2019, p. 375). Os resultados das variáveis semânticas confirmam que esse tipo de variável não desempenha papel significativo na inserção do artigo definido, pois todos os fatores, exceto o favorecimento exercido pela *referência ao SN do documentador*, apresentam valores neutros quanto à inserção do artigo definido, conclui Figueiredo (2019, p. 376).

No grupo *número no SN*, o SN plural se mostra favorável à inserção do artigo e o SN singular com semântica plural favorece fortemente o determinante \emptyset , enquanto o SN singular revela um peso neutro. Infere-se, então, que “a semântica plural do nome, por si só, não é suficiente para determinar a inserção” (FIGUEIREDO, 2019, p. 377). A aquisição do português-L2 pelos falantes do grupo níger-congolês pode estar na base dos resultados apresentados, aponta Figueiredo (2019, p. 378): na evolução da aquisição do artigo, (i) os falantes inicialmente não distinguem morfologicamente o singular e o plural; (ii) em seguida, aplicam o artigo no singular; (iii) inserem a marca formal de plural, inicialmente no determinante, e, (iv) por fim, nos outros elementos do SD, com variação na realização da concordância (FIGUEIREDO, 2010, p. 207)³⁶. A variável *tipo de artigo definido*, cruzada com o *número do SN* e *faixa etária*, confirma as tendências de que (i) os SNs singulares favorecem os SNs nus, enquanto os plurais favorecem a inserção do artigo definido; (ii) os almosarifanos utilizam mais o artigo masculino do que o feminino, devido ao valor de *default* encontrado no artigo masculino e no pronome pessoal com função de determinante, cuja utilização é encontrada apenas nas faixas 2 e 3, extinguindo-se na faixa 1 (FIGUEIREDO, 2019, p. 381-382). Em suma, os fatores sócio-estruturais se mostram mais relevantes do que os semânticos no condicionamento da variação do artigo definido do PA (FIGUEIREDO, 2019, p. 382).

2.4. PESQUISAS RECENTES SOBRE SDs NUS DE REFERÊNCIA DEFINIDA EM LÍNGUAS CRIOULAS DE SÃO TOMÉ

O Crioulo Caboverdiano (CCV), investigado por Baptista (2007), e o Santome (ST), descrito por Alexandre e Hagemeijer (2007), são duas línguas crioulas de base lexical portuguesa faladas em São Tomé. Ambas apresentam aspectos semânticos e sintáticos nos SNs nus e/ou

³⁶ apud Figueiredo (2019, p. 378).

definidos que ocorrem tanto nas variedades do português apresentadas na seção 2.3 como no português de Monte Café, foco desta pesquisa.

2.4.1 Os SDs nus no Crioulo Caboverdiano (BAPTISTA, 2007)

Baptista (2007) analisa no CCV aspectos semânticos e estruturais dos SNs nus, bem como estratégias de marcação de definitude no singular e no plural. No domínio da definitude do CCV, os demonstrativos *kel* (sg.) e *kes* (plur.), além de exercerem a sua função principal, também podem assumir o papel de determinante definido. Baptista (2007, p. 63) ressalta que é comum nas línguas em geral, esse caminho evolutivo: os artigos definidos tendem a se originar dos demonstrativos, e os artigos indefinidos (no CCV: *un*, *uns*) geralmente derivam do numeral singular. Embora os determinantes fonéticos sirvam para identificar SNs definidos e indefinidos, eles podem ser substituídos pelo determinante \emptyset se houver informações semântico-sintáticas suficientes sobre a (in)definitude do nome na oração (BAPTISTA, 2007, p. 70).

Os nomes nus são muito produtivos no CCV e podem ter várias interpretações³⁷, das quais nos interessam as [+ específicas] e [+ definidas], singular e plural (BAPTISTA, 2007, p. 70-72):

86) Plural [+ específico][+ definido]:

Kaza di es aldeia e baratu.

*As casas dessa aldeia são baratas.*³⁸

87) Singular [+ específico][+ definido]:

Ta ben txeu djenti ta ben buska-m pa'N konta-l storia di kel organizason.

Muitas pessoas vêm me procurar para que eu conte a elas *a história da organização*.

Baptista (2007, p. 75) descreve, em seguida, algumas condições sob as quais os nomes nus [+ específicos][+ definidos] são realizados:

³⁷ Os SNs nus podem ter interpretação (i) genérica, (ii) plural específica/definida, (iii) plural específica/ indefinida, (iv) plural não-específica/não-definida, (v) singular específica/indefinida, (vi) singular não-específica/não-definida, (vii) singular definida/não-específica, (viii) singular específica/definida (BAPTISTA, 2007, p. 70-71).

³⁸ A tradução dos dois exemplos é nossa.

88) Quando a singularidade da entidade a torna de conhecimento geral, o que se aplica a fenômenos naturais, como a ‘chuva’, ou entidades únicas, como o ‘sol’:

Sol ta ben mas txeu, ta da kumida mas txeu.

O sol sai mais, há mais comida.

89) Quando o elemento faz parte da experiência comum do falante e do ouvinte:

N ben ta ntende kuse ki *xeфри* ta konbersa-m.

Aos poucos, fui entendendo sobre o que o *chefe* está falando comigo.

90) Quando a primeira menção a um substantivo é feita com um determinante fonético, a segunda é tipicamente com nomes nus³⁹:

Es gosta di Vieira na Praia, es ka xa-l bai, el fi ka k’ un mudjer la. [...] *Mudjer* gosta d’ el, fi ka ku el ala.

Gostavam do Vieira na Praia, não o deixavam ir, ele ficava com uma mulher lá. [...] *A mulher* gostava dele e ficava com ele lá.

A respeito da interpretação dos nomes nus, Baptista (2007, p. 76) afirma que Lucchesi (1993) observou corretamente que nomes nus no CCV podem ter leituras (i) específica e definida, (ii) indefinida e específica ou (iii) indefinida e não-específica. Essa generalização, porém, não se aplica da mesma forma aos SNs nus singulares e plurais: SNs nus singulares com semântica plural podem ser interpretados das três formas, tanto na posição de sujeito quanto na de objeto (BAPTISTA, 2007, p. 76); por outro lado, a posição sintática ocupada pelos SNs nus singulares interfere na interpretação que recebem: os SNs nus singulares definidos (informações conhecidas) geralmente ocupam a posições de sujeito, e nomes nus singulares indefinidos (informações novas) tendem a aparecer em posições de objeto (BAPTISTA, 2007, p. 76-77).

Outro aspecto do CCV que Baptista (2007, p. 82-84) descreve é a marcação de número nos SNs definidos, que tem três estratégias de pluralização:

91) Determinante quantificador⁴⁰ e nome morfológicamente singular:

³⁹ Baptista (2007, p. 75) afirma que essa constatação foi inicialmente feita por Lucchesi (1993, p. 81).

⁴⁰ Demonstrativo, possessivo, numeral ou quantificador flutuante.

N ten *seti* fidju, *sinku* matxu, *dos* femia.

Eu tenho *sete* filhos, *cinco* meninos, *duas* meninas.

92) Sufixação do -s nos nomes nus [+ animados] [+ definidos]:

Ta bende un balai di pexi, nu ta kre pa kria *fidjus*.

‘Nós vendemos uma cesta de peixes, precisamos disso para criar *os filhos*.’

93) Determinante e nome [+ humano] pluralizados:

Kes djentis, bedju, txeu ka ta baba skola p’ e djuda mamai kiria mininu.

‘*Essas pessoas*, antigamente, não iam à escola para ajudar suas mães com os filhos.’

A forma de pluralização mais comum é marcação de plural através de um quantificador, que acompanha um nome morfológicamente singular (exemplo 91); os nomes [+ animados][+ definidos] podem apresentar marcação morfológica de plural (exemplo 92); e, em casos raros, o marcador plural -s aparecem tanto no determinante quanto no nome (exemplo 93), estratégia exclusiva para os nomes plurais [+ humanos]. Assim, a *definitude* e a *animacidade* são, como argumenta Baptista (2002, 2003)⁴¹, os dois fatores que mais influenciam na sufixação do plural no CCV (BAPTISTA, 2007, p. 84).

2.4.2 Os SDs nus no Santome (ALEXANDRE E HAGEMEIJER, 2007)

Alexandre e Hagemeyer (2007) analisam, nos níveis de interpretação da frase e do discurso, as propriedades dos SNs nus do ST, um dos quatro crioulos de base portuguesa do Golfo da Guiné. O *corpus* utilizado para esse estudo é composto por mais de 20 horas de fala gravada espontânea, que correspondem sobretudo a histórias folclóricas, e de fala induzida (ALEXANDRE; HAGEMEIJER, 2007, p. 39). Os marcadores que atribuem (in)definitude e número no ST são *ũa* ‘um(a)’ [- definido][- plural], e *inen* ‘os/as’ [+ definido][+ plural]. Os SNs nus atuam, então, como suas contrapartidas, expressando as referências [+ plural][- definida] e [- singular][+ definida]. (ALEXANDRE; HAGEMEIJER, 2007, p. 39-40).

⁴¹ apud Baptista (2007, p. 84).

A realização de *inen + nome*, sem a presença de outros especificadores, sempre ocorre com substantivos de referência [+ humana], como pode-se observar no exemplo (94)⁴², dado por Alexandre e Hagemeijer (2007, p. 41):

94) *Inen funxônariu ska bi golo pixi blatu ô.*

*Os funcionários estão procurando peixe barato.*⁴³

A preferência do *inen* por substantivos [+ humanos] pode ser considerada “uma consequência de sua sensibilidade a um princípio semântico que considera [+ humanos] como sendo mais individualizado do que itens que são [- humanos]”⁴⁴ (ALEXANDRE; HAGEMEIJER, 2007, p. 41). Outro fato que argumenta a favor da leitura definida de *inen* é que esse determinante não permite interpretações genéricas. Através de dois exemplos, os autores ilustram o bloqueio causado pelo *inen*: a única interpretação possível para *inem (ngê)* ‘as pessoas’ no exemplo (95) é que se trata de *um grupo de pessoas familiar para o falante*, não sendo possível se referir a *toda e qualquer pessoa de São Tomé* – essa interpretação se torna plausível no exemplo (96), em que *ngê* ‘pessoas’ constitui um SN nu [+ plural][- definido] (ALEXANDRE; HAGEMEIJER, 2007, p. 41).

95) *Inen (ngê) santome sêbê kanta.*

As pessoas de São Tomé cantam bem.

96) *Ngê santome sêbê kanta.*

Pessoas de São Tomé cantam bem.

Na maior parte do *corpus*, porém, são encontrados SNs modificados simultaneamente por *inen* e por outros constituintes que indicam especificidade, como orações relativas, adjetivos ou possessivos – que ocorrem em posição pós-nominal no ST. Uma estrutura particularmente frequente, nesse caso, é o uso do *inen* e do *se* (marcador de especificidade padrão) no SN, como Alexandre e Hagemeijer (2007, p. 41) mostram a partir dos exemplos (97) e (98). Esse maior grau

⁴² Os exemplos de Alexandre e Hagemeijer (2007) foram traduzidos para o inglês pelos autores, e a tradução para o português de todos os exemplos desse texto é nossa.

⁴³ Tradução nossa.

⁴⁴ “The plural marker shows a preference for nouns with the feature [+human], which we consider a consequence of its sensitivity to a semantic principle that takes [+human] as being more individuated than items that are [-human].” Tradução nossa.

de referencialidade, devido à coocorrência dos marcadores de definitude e especificidade, é muito comum com nomes [- humanos], que não são encontrados, no *corpus*, acompanhados exclusivamente do *inen*, sem outros marcadores adicionais (ALEXANDRE; HAGEMEIJER, 2007, p. 41).

97) *So inen ja se so sun Fenanji na ska txila vin fa.*

Então *esses dias* Sr. Fenanji não está extraindo vinho da palma.

98) *Kê kwa ku bô mêsê pa a fe inen manu se bô?*

O que você quer que eles façam com *seus irmãos*?

O marcador *se* ‘o, este, estes’, como atribuidor de especificidade, refere-se a objetos cujo conhecimento é compartilhado entre os interlocutores. Alexandre e Hagemeijer (2007, p. 44) alegam que o *se* não deve ser considerado um artigo definido, devido à possibilidade de coocorrer com os marcadores *inen* e *ũa*; além disso, os SNs nus geralmente possuem leitura definida.

Apesar de o SN nu definidos ter uma gama de possibilidades interpretativas, os autores reforçam que ele precisa estar previamente presente no discurso para se tornar familiar ao ouvinte (ALEXANDRE; HAGEMEIJER, 2007, p. 49). Os SNs nus têm algumas particularidades relacionadas à animacidade e ao número: (i) as formas respeitosas de se referir a um homem ou uma mulher, *sungê* e *sangê*, são interpretadas exclusivamente como singulares, como em *Sungê koplá pixi* ‘O senhor compra peixe’; (ii) embora outros SNs humanos possam receber interpretação singular ou plural, eles aparecem majoritariamente como singulares no *corpus*; (iii) os nomes nus [-humanos] têm geralmente semântica plural (ALEXANDRE; HAGEMEIJER, 2007, p. 49). Acrescentando, ainda, informações estruturais, Alexandre e Hagemeijer (2007, p. 49-50) afirmam que os SNs nus na posição de sujeito, associada a informações já conhecidas pelos interlocutores, diferem daqueles realizados como objeto, que geralmente contém informações novas. Sempre que os SNs nus apresentam uma informação nova na posição de objeto, têm uma leitura plural indefinida, exceto nos casos especiais de *sangê* e *sungê*, mas quando estão previamente presentes de alguma forma, podem receber leitura definida – singular ou plural (ALEXANDRE; HAGEMEIJER, 2007, p. 50).

Partindo do pressuposto de que o determinante zero é mais primitivo do que determinantes fonéticos, é esperado, de acordo com Alexandre e Hagemeijer (2007, p. 50), que os SNs nus sejam,

inicialmente, associados a leituras definidas singulares na posição de sujeito, e a leituras indefinidas plurais, na posição de objeto. As línguas que não desenvolveram completamente um sistema de determinantes definidos emprestariam, assim, interpretações definidas aos SNs nus, ao passo que as línguas cujo sistema de determinante definido é completo podem começar a rejeitar a leitura definida de SNs nus. A partir disso, os autores sugerem que o sistema determinante em ST se encontra em desenvolvimento (ALEXANDRE; HAGEMEIJER, 2007, p. 50).

2.5 SÍNTESE

Este capítulo apresentou a fundamentação teórica da presente dissertação. A partir do trabalho Lyons (1999), a noção semântica de *definitude* foi descrita a partir da sua relação com outros conceitos basilares para o seu entendimento: familiaridade, especificidade, contabilidade e número animacidade e tipo de posse. O autor sugere, ainda, que a definitude também corresponde a uma categoria gramatical existente na sintaxe de algumas línguas, expressa pelo núcleo funcional D, que marca formalmente o que denomina como a noção semântica/pragmática de definitude. Em seguida, foi apresentado o Parâmetro do Mapeamento Nominal de Chierchia (1998), uma proposta teórica sobre a tipologia das línguas naturais que as classifica de acordo com o que os seus SNs podem denotar: nomes de tipo e/ou predicados. Posteriormente, foram comentadas as evidências demonstradas por Schmitt e Munn (1999) de que o PB não se enquadra no Parâmetro de Chierchia (1998). Posteriormente, foram apresentados estudos que analisam os nomes nus de referência definida em três variedades do português e em dois crioulos falados em São Tomé: o PB (WALL, 2013), o PABH (BAXTER; LOPES, 2009; NEVES, 2019), o PA (FIGUEIREDO, 2019), o CCV (BAPTISTA, 2007) e ST (ALEXANDRE; HAGEMEIJER, 2007). Os textos mencionados nos oferecem, assim, uma base consistente para a formulação de variáveis linguísticas e extralinguísticas que contêm possíveis fatores condicionantes para o uso dos SDs nus referenciais no Português de Monte Café (PMC).

3. A COMUNIDADE, OS DADOS E A METODOLOGIA

O presente capítulo descreve a comunidade linguística focada, a natureza do *corpus* e a metodologia utilizada para efetuar a análise da variação no uso do determinante de referência definida. Inicia-se a discussão com a apresentação das peculiaridades do contexto sócio-histórico e linguístico da comunidade na seção 3.1. Em seguida, as seções 3.2 e 3.3 dedicam-se, respectivamente, à descrição do caráter do *corpus* e à definição e explicação da variável dependente e das variáveis independentes que instrumentalizam as hipóteses sobre o condicionamento da

variação em foco. Finalmente, a seção 3.4 explica os procedimentos de codificação e quantificação seguidos na análise estatística dos dados e a seção 3.5 sintetiza o que foi apresentado.

3.1 A AQUISIÇÃO DO PORTUGUÊS NA GRANDE ROÇA DE MONTE CAFÉ

A história da ilha de São Tomé e Príncipe é dividida em dois momentos, o da habitação (1493-1520) e o da plantação de cana-de-açúcar (1520-1822), que resultaram no surgimento de quatro línguas crioulas de base lexical portuguesa: o Santome, o Angolar, o Lung'ie e o Fa d'Ambô (HAGEMEIJER, 2009, p. 2). Essas línguas surgiram durante a escravização de africanos trazidos, inicialmente, do antigo Reino do Benim, e em seguida, das zonas Bantu (primeiro do Congo, e depois, da Angola), quando houve uma necessidade imediata de comunicação com os colonos portugueses (HAGEMEIJER, 2009, p. 2-4). Mudanças socioeconômicas e demográficas caracterizaram o ciclo do cacau e do café, e, no início da década de 1850, prevendo a libertação dos estrangeiros escravizados (concretizada em 1876), alguns fazendeiros começaram a importar e contratar mão de obra angolana (NASCIMENTO; GOMES DIAS, 1989)⁴⁵. Em 1854, ano da libertação dos escravizados nativos⁴⁶ (denominados *forros* a partir de então), surgiu a roça de Monte Café (BAXTER, 2002, 2004), cuja localização pode ser vista no Mapa 1.

⁴⁵ apud Baxter (2004, p. 100).

⁴⁶ Neves (1929, p. 55) apud Baxter (2002, p. 8).

Mapa 1 - São Tomé e Príncipe



Fonte: Mapa base do National Geographic's MapMaker Interactive (editado)

A grande roça produtora de cacau e café, considerada uma das mais importantes de São Tomé, pertencia aos portugueses e está localizada no distrito de Mé-Zóchi. Na segunda metade do século XIX, quando cada roça era conhecida pela língua africana mais usada, em Monte Café predominava o umbundu, falado pelos angolanos e seus descendentes, os *tongas* (ROUGÉ, 1992)⁴⁷. Os angolanos corresponderam, durante todo o período de contratação, à maior parte da mão de obra estrangeira de Monte Café; também foram contratados, em menor proporção, moçambicanos e caboverdianos. Sobre esse período, Cadbury (1969)⁴⁸ traz duas informações relevantes: enquanto os trabalhadores estrangeiros foram repatriados antes de 1909, os angolanos só conseguiram sair da ilha no século XX; e mesmo quando esses angolanos estavam voltando à sua terra de origem, os *tongas*, nascidos em São Tomé, foram obrigados a permanecer na ilha. Os funcionários viviam, em grande parte, confinados nas roças até meados de 1940 (BAXTER, 2002, p. 10).

⁴⁷ apud Baxter (2002, p. 10).

⁴⁸ apud Baxter (2002, p. 9).

Dentro das roças, os trabalhadores vivenciavam uma situação em que, embora algumas de suas línguas de origem funcionassem como língua franca, o português era a língua de comunicação entre eles e os seus supervisores (que eram, em sua maioria, europeus⁴⁹). O português foi aprendido por eles em situação de Transmissão Linguística Irregular (BAXTER, 1995; LUCCHESI; BAXTER, 2009) em um conjunto peculiar de contextos de aquisição. Os dados linguísticos primários (DLP) para a aquisição e transmissão linguística eram muito diversos. Baxter (2002, p. 11-14) afirma que, por um lado, houve crianças de pais africanos, cuja L1 era a língua africana dos pais, que aprendiam o português como L2 através do contato com outros adultos e crianças, principalmente no campo. Por outro lado, houve crianças de casais tonga + africano(a), e de casais tongas, que naturalmente teriam adquirido o umbundo juntamente com o português L1 falado pelos pais, mas ao mesmo tempo haveria outros modelos linguísticos fornecidos por outros serviçais. Em ambos os contextos, a aquisição do português e do umbundo acontecia no campo, no quintal e na senzala (BAXTER, 2002, p. 11-14). O regime contratual contínuo tornou esse contexto ainda mais complexo já que, em intervalos de 3-5 anos, novos trabalhadores eram contratados, e, conseqüentemente, crescia o número de adultos falantes de um estágio inicial de português-L2.

3.2 A CONSTITUIÇÃO DO *CORPUS* TONGA

A fim de analisar a variação no uso do determinante nos SDs sujeitos e objetos diretos [+específicos][+ definidos] no desempenho linguístico de falantes das primeiras gerações que tiveram o português como língua nativa em Monte Café, foram estudadas 18 entrevistas sociolinguísticas do *Corpus* Tonga⁵⁰ (BAXTER, 1998-2000), gravadas em 1998. As entrevistas estão distribuídas uniformemente entre homens e mulheres das faixas etárias 1 (de 20 a 40 anos), 2 (41 a 60 anos) e 3 (acima de 60 anos). Os informantes têm, em geral, o mesmo perfil social: são analfabetos (sobretudo na faixa 3) ou semianalfabetos (estudaram até a escola primária⁵¹); têm contato muito limitado com o português na mídia (ex. televisão) e têm pouco acesso à comunidade fora da grande

⁴⁹ Cadbury (1969 [1910]) apud Baxter (2002, p. 10).

⁵⁰ Parte do projeto *Testing the hypothesis of prior creolization* patrocinado pelo *Australian Research Council*, o *corpus* é constituído por 26 entrevistas de 1h30min. Uma cópia deste *corpus* está arquivado no acervo do Projeto Vertentes da Universidade Federal da Bahia.

⁵¹ O ensino escolar santomense é distribuído em básico, da 1ª à 6ª classe, e secundário, da 7ª à 12ª classe (HAGEMEIJER; GONÇALVES; AFONSO, 2018, p. 69). Os informantes que afirmaram ter ido à escola relatam que estudaram até a 3ª, 4ª ou 6ª classe.

roça. A fim de elucidar o cenário de aquisição linguística desses informantes, o Quadro 1 os lista conforme faixa etária, parentesco e ida à escola:

Quadro 1 - Principais fontes dos DLP dos informantes de Monte Café

Faixa etária	Informante	Principal fonte dos DLP	Ida à escola
I	[MCJCAH1]	Pais tongas	Até a 6ª classe
	[MCFJBH1]	Mãe angolana	Até a 6ª classe
	[MCFACH1]	Pais tongas	Até a 4ª classe
	[MCANJM1]	Pais tongas	Até a 4ª classe
	[MCJNKM1]	Pais tongas	Foi à escola
	[MCLVLM1]	Pais angolanos	Até a 6ª classe
II	[MCJCDH2]	Mãe tonga e pai angolano	Não menciona
	[MCALEH2]	Pais tongas	Não foi à escola
	[MCAJFH2]	Pais tongas	Até a 4ª classe
	[MCISMM2]	Pais angolanos	Não foi à escola
	[MCMFNM2]	Pais tongas	Não foi à escola
	[MCMLOM2]	Pais tongas	Foi por 3 anos
III	[MCMCGH3]	Pais tongas	Não foi à escola
	[MCAJHH3]	Pais tongas	Não foi à escola
	[MCPFIH3]	Pais tongas	Não foi à escola
	[MCSGPM3]	Pais tongas	Não foi à escola
	[MCCEQM3]	Pais angolanos	Foi à escola
	[MCINRM3]	Pais tongas	Não foi à escola

Pode-se inferir, assim, que a maioria dos informantes aprenderam o português-L1 dos pais tongas, enquanto os filhos de angolanos adquiriram como L1 o português-L2 dos pais. Junto com o português, todos têm algum conhecimento do *umbundu*, mas suas competências variam⁵²: enquanto os informantes das faixas 3 e 2 são falantes dessa língua, na maior parte da faixa 1 o

⁵² Para mais detalhes sobre as diferentes situações de bilinguismo desses falantes, consultar quadro de *Perfil aquisicional dos informantes tongas* (SANTANA, 2019, p. 59-60).

conhecimento é passivo (compreendem, mas não a utilizam ativamente). Quanto a outras línguas com que afirmam terem tido contato, alguns entrevistados mencionam o *forro*, o *caboverdiano*, o *kimbundo*, a “língua de tonga”, a “língua de Angola” e a “língua de Moçambique”.

3.3 A VARIÁVEL DEPENDENTE E AS VARIÁVEIS INDEPENDENTES

Observações preliminares da variação no uso dos determinantes [+ específicos][+ definidos] na posição de sujeito e objeto direto no português de Monte Café possibilitaram a formulação de uma lista de hipóteses de trabalho sobre fatores morfossintáticos, semânticos e extralinguísticos que poderiam orientar esse uso. Assim, organizou-se uma chave de codificação para classificar cada determinante de acordo com as sugestões de potenciais fatores condicionantes.

3.3.1. A variável dependente

A análise dos determinantes usados nos SDs [+ específicos][+ definidos] no português de Monte Café se constituiu a partir de três possibilidades de uso do D⁵³:

(i) Determinante Ø.

- (1) **Guera** veio assi memo, nossos preto é que ranjô isso. [MCPFIH3]
- (2) Só, meu avô pediu no mato... depois foram apanhar **corpo** só... [MCLVLM1]

(ii) Artigo definido.

- (3) **A guera** gitô muito o pais. [MCPFIH3]
- (4) Sexta-feira é que gente recebeu **os pagamento** agora, gente recebeu, recebi quarenta e sete conto que eu recebi. [MCAJHH3]

(iii) Demonstrativo.

- (5) Agora **isso hospitare** no te dêru, no cura, no da você no da nada, nada, nada. [MCCEQM3]

⁵³ No *corpus*, o demonstrativo é observado na função de representante de referência definida sem valor dêitico.

(6) ...vai pa cima, trabaia bocadinho, ganha dinheiro pa stenta **issos garoto**, quatro conto isso tempo é quê? [MCMLOM2]

Nos estudos referentes ao sistema de determinantes de variedades do português, alguns pesquisadores tratam a variação na representação do SD como uma variável binária (o artigo definido e a sua ausência) (BAXTER; LOPES, 2009; FIGUEIREDO, 2019), enquanto outros a trataram como uma variável trinomial (NEVES, 2019). No presente estudo, o demonstrativo e o artigo definido foram reunidos em uma variante, e a variável dependente foi tratada como binária, formada pelo determinante zero e pelos determinantes foneticamente realizados (artigo definido + demonstrativo).

3.3.2 As variáveis independentes linguísticas

A investigação de trabalhos prévios sobre o uso dos determinantes em variedades do português brasileiro (WALL, 2013), de Helvécia (BAXTER; LOPES, 2009; NEVES, 2019) e de Almojarife (FIGUEIREDO, 2019), bem como no Santome (ALEXANDRE E HAGEMEIJER, 2007) e no Crioulo Cabo-Verdiano (BAPTISTA, 2007), revelou uma série de variáveis que condicionam a realização do determinante zero. No presente estudo, para finalidades comparatistas, optou-se por avaliar a validade dessas mesmas variáveis. O Quadro 2 apresenta as variáveis linguísticas independentes, cujas hipóteses serão descritas a seguir.

Quadro 2 - Variáveis independentes linguísticas

VARIÁVEIS LINGUÍSTICAS	
Função sintática do SD	- Sujeito - Objeto direto
Posição do SD na oração	- Sujeito - Sujeito em posição marcada - Voz passiva - Objeto direto - Foco - Existencial

Animacidade do substantivo⁵⁴	- Animado - Inanimado
Familiaridade	- Anafórica - Referência cruzada - Catafórica - Situacional - Conhecimento geral
Nomes contáveis/ não-contáveis	- Contável - Não-contável
Número no SD	- Singular - Plural - Singular com leitura de plural - Singular com possibilidade de plural
Tipo de posse	- Inalienável - Alienável
Outro constituinte no SD	- Oração relativa - Adjetivo pré-nominal - Adjetivo pós-nominal - Sintagma preposicionado - Possessivo pré-nominal - Possessivo pós-nominal - Advérbio locativo - Nenhum outro modificador

1) **Função sintática do SD:** a partir dessa variável, busca-se avaliar se as funções sintáticas de sujeito e OD condicionam o uso do determinante Ø nos SDs [+ específicos][+ definidos]. No PABH (BAXTER; LOPES, 2009, p. 322), o adjunto é o único fator que favoreceu o aparecimento do artigo definido, e o sujeito e o OD são as funções sintáticas que menos inibiram o seu uso; no PA (FIGUEIREDO, 2019, p. 365), esse fator não é considerado estatisticamente relevante; no CCV (BAPTISTA, 2007, p. 77), os SDs nus singulares tendem a ser definidos na posição de sujeito. Supõe-se, assim, que a posição de sujeito deve favorecer o uso do determinante Ø.

⁵⁴ Durante a coleta de dados, encontramos 3 ocorrências referentes ao *diabo*, todas com determinante zero, o que levou à criação da variante **sobrenatural** nesse grupo. A variante foi posteriormente descartada, pois não houve outras ocorrências.

2) **Posição do SD na oração:** essa variável é proposta no intuito de verificar se o uso do SD em posição não-marcada⁵⁵ ou marcada⁵⁶ condiciona o uso do determinante zero. O fator foi inicialmente inserido no estudo dos OD [+ específicos][+ definidos] do PABH (NEVES, 2019, p. 158), mas posteriormente retirado da análise. No PB, observa-se que os SDs nus [+ definidos][+ específicos] ocorrem predominantemente nos sujeitos e objetos em posição pré-verbal (WALL, 2013, p. 237-242). Esse resultado aponta para um possível favorecimento do determinante \emptyset nos sujeitos em posição não-marcada.

3) **Animacidade do substantivo:** o traço [+ humano], no PA (FIGUEIREDO, 2019), e o [+ animado], no CCV (BAPTISTA, 2007), favorecem o determinante zero, correspondendo, assim, à relação proposta por Lyons (1999) entre animacidade e definitude: o traço [+ humano], por ser mais saliente, favorece o determinante \emptyset . No PABH (BAXTER; LOPES, 2009; NEVES, 2019), essa variável não foi considerada estatisticamente relevante. No ST, o *inen* (marcador de definitude e plural) mostra uma preferência por nomes [+ humanos]; quando ocorre com [- humanos], costuma requerer mais de um especificador (ALEXANDRE; HAGEMEIJER, 2007, p. 40-41). Com base nos estudos apresentados, o [+ humano] deve favorecer o determinante \emptyset nos dados de Monte Café.

4) **Familiaridade:** a relação proposta por Lyons (1999) entre a definitude e a familiaridade, i.e., o modo como o ouvinte identifica o objeto na situação de fala, se mostrou relevante no PABH, no PB e no CCV. No CCV, o conhecimento prévio e a anáfora propiciam o SD nu (BAPTISTA, 2007, p. 74); no PB, a anáfora e a referência cruzada favorecem o determinante \emptyset (WALL, 2013, p. 237-238); nos ODs [+ específicos][+ definidos] do PABH, a catáfora e a referência cruzada favorecem o determinante \emptyset (NEVES, 2019, p. 180-183). Com essa variável, busca-se avaliar a hipótese de que a maneira como o ouvinte identifica os SDs condiciona o uso do determinante zero.

5) **Nomes contáveis/não-contáveis:** já foi mostrado em estudos de diferentes variedades do português (BAXTER; LOPES, 2009; FIGUEIREDO, 2019) que a proposta de Chierchia (1998)

⁵⁵ Sujeito antes do verbo e OD depois do verbo.

⁵⁶ Sujeito depois do verbo e OD antes do verbo.

de que os substantivos contáveis nus se referem a *tipos* não é regra na nossa língua. Tanto os nomes [+ contáveis] quanto os [- contáveis] podem ser [+ específicos] com o determinante Ø. Embora esse fator não seja considerado estatisticamente relevante nos estudos mencionados, observa-se no PABH um certo favorecimento dos SDs nus nos nomes [- contáveis] (NEVES, 2019, p. 198). Espera-se que os substantivos [- contáveis] favoreçam o determinante Ø nos SDs [+ específicos][+ definidos] do PMC.

6) **Número no SD:** nos estudos do uso dos determinantes no português (BAXTER; LOPES, 2009; NEVES, 2019; FIGUEIREDO, 2019), os SDs singulares favorecem o determinante zero, e o oposto acontece com os SDs plurais. No ST, há uma partícula (*inen*) que indica definitude e plural, contrastando numericamente com os SDs nus – singulares definidos (ALEXANDRE; HAGEMEIJER, 2007). A fim de investigar se os singulares também favorecem o determinante Ø e se a marcação de plural condiciona a forma do determinante, esse fator foi adicionado.

7) **Tipo de posse:** a relação entre a definitude e as posses alienável e inalienável⁵⁷ é proposta por Lyons (1999). Ele afirma que, em várias línguas, a posse inalienável é marcada com uma forma morfológicamente mais simples (LYONS, 1999, p. 128). No PABH, esse fator não foi considerado estatisticamente relevante (NEVES, 2019, p. 187-189). A inclusão dessa variável possibilitará a verificação de se a posse inalienável favorece o uso do determinante no PMC.

8) **Outro constituinte no SD:** no CCV (BAPTISTA, 2007) e no PABH (NEVES, 2019, p. 274-280), a presença de outros constituintes no SD favorece o SD nu, mas o comportamento de cada tipo de constituinte varia: os possessivos e os sintagmas preposicionados favorecem determinante Ø no PABH, enquanto a oração relativa desfavorece (BAXTER; LOPES, 2009; NEVES, 2019). Supõe-se que a presença de outros componentes no SD favorece o determinante Ø no PMC.

3.3.3 As variáveis independentes sociais

⁵⁷ Lyons (1999) aponta que a diferença entre uma posse alienável e uma inalienável é que, enquanto a primeira pode ser transferida, vendida ou cedida, a segunda está ligada ao possuidor de forma intrínseca, como *familiares* e *partes do corpo*.

A escolha das variáveis sociais (no quadro 3 a seguir) decorre da constituição sócio-histórica da comunidade.

Quadro 3 - Variáveis independentes sociais

VARIÁVEIS SOCIAIS	
Faixa etária	- Faixa 1 (20 a 40 anos) - Faixa 2 (41 a 60 anos) - Faixa 3 (acima de 60 anos)
Sexo	- Homem - Mulher

1) **Faixa etária:** um dos objetivos deste estudo é observar o desenvolvimento do sintagma determinante em posição de sujeito e objeto com leitura [+Específica] [+Definida] do PMC através da diacronia em tempo aparente. Nos estudos analisados, a variação mostrou um potencial perfil de aquisição do artigo definido (BAXTER; LOPES, 2009; NEVES, 2019; FIGUEIREDO, 2019). Com base nos resultados analisados e no contexto sócio-histórico de Monte Café, o mesmo perfil aquisicional deve ser encontrado na fala dessa comunidade.

2) **Sexo:** o contraste social entre os homens e mulheres é um dos mais claros e consistentes, segundo Labov (1990, p. 205 apud TAGLIAMONTE, 2012, p. 32) e a variável sexo está presente nos estudos sociolinguísticos. No PABH, esse fator não é estatisticamente significativo (NEVES, 2019, p. 189-194). No PA, a fala dos homens desfavorecem o uso do SD nu (FIGUEIREDO, 2019, p. 371), e no estudo dos OI do PMC (SANTANA, 2019, p. 88), os homens desfavorecem a forma não-padrão. Com base nos resultados apresentados, espera-se que a fala das mulheres favoreça, e a dos homens desfavoreça, o determinante Ø no PMC.

3.4 O TRATAMENTO DOS DADOS

A pesquisa desenvolvida seguiu as seguintes etapas: (i) delimitação das variáveis dependente e independentes; (ii) levantamento de dados a partir do *corpus*; (iii) codificação dos dados com em uma tabela no Excel; (iv) processamento dos dados no GOLDVARB-X (SANKOFF, TAGLIAMONTE; SMITH, 2005) e (v) análise e interpretação dos resultados.

Os dados foram codificados no Excel porque uma codificação em células oferece maior segurança no sentido de minimizar a margem de erro na codificação e, além disso, o Excel permite sortear e listar ocorrências de variáveis individuais. A fim de avaliar o condicionamento da variação, foi utilizado o pacote do GOLDVARB-X (SANKOFF, TAGLIAMONTE; SMITH, 2005).

Foram identificados e codificados todos os SDs [+ específicos][+ definidos] em posição de sujeito e objeto direto nas entrevistas. Durante o levantamento de dados, foram eliminadas ocorrências que continham encontro de palavras terminadas e/ou começadas em *a* e *o*, como em *Daqui uma hora de relos ou duas pronto, já não sinto nada a mão fica toda vazia.* [MCAJHH3], em que não se pode garantir que o artigo foi foneticamente realizado. Posteriormente, foram eliminados todos os fatores que não apresentavam o elenco completo de variantes (ART, DEM, ZERO). Após os ajustes, foram encontrados 799 dados no total, entre SDs com determinante zero, artigo definido e demonstrativo.

3.5 SÍNTESE

Neste capítulo, foi brevemente apresentado o contexto sócio-histórico de Monte Café, bem como as condições sob as quais os tongas viviam e os diferentes dados linguísticos primários (DLP) aos quais foram expostos durante a aquisição por Transmissão Linguística Irregular (BAXTER, 1995; LUCCHESI; BAXTER, 2009). Em seguida, caracterizou-se o *corpus* utilizado nesta pesquisa, considerando a principal fonte de DLP dos falantes, suas faixas etárias e ida à escola. Por fim, foram descritos os aspectos relacionados à metodologia aplicada no presente trabalho. No capítulo seguinte, serão apresentados os resultados estatísticos gerados pelo programa GOLDVARB-X, a partir dos quais se pautam as interpretações sobre o encaixamento linguístico e social de cada variável dependente e sua relação com a aquisição em situação de contato linguístico.

4. RESULTADOS

O presente capítulo expõe os resultados obtidos através da análise quantitativa dos dados da variável dependente em estudo, o tipo de determinante dos SDs [+ específicos] [+ definidos] em posição de sujeito e objeto direto, cujas variantes são o determinante zero, o artigo definido e o demonstrativo. O capítulo está dividido em quatro seções. Na seção 4.1, são apresentadas a distribuição geral dos SDs e são elencadas as variáveis selecionadas (em ordem de seleção) e rejeitadas (em ordem de rejeição) pelo GOLDVARB-X. À continuação, na seção 4.2, são apresentadas e comentadas as variáveis independentes selecionadas, e, na seção 4.3, as consideradas estatisticamente não significativas para a realização do determinante zero em Monte Café. Por fim, a seção 4.4 sintetiza os resultados apresentados.

4.1 A DISTRIBUIÇÃO GERAL DE SDs REFERENCIAIS DEFINIDOS NAS POSIÇÕES DE SUJEITO E OBJETO

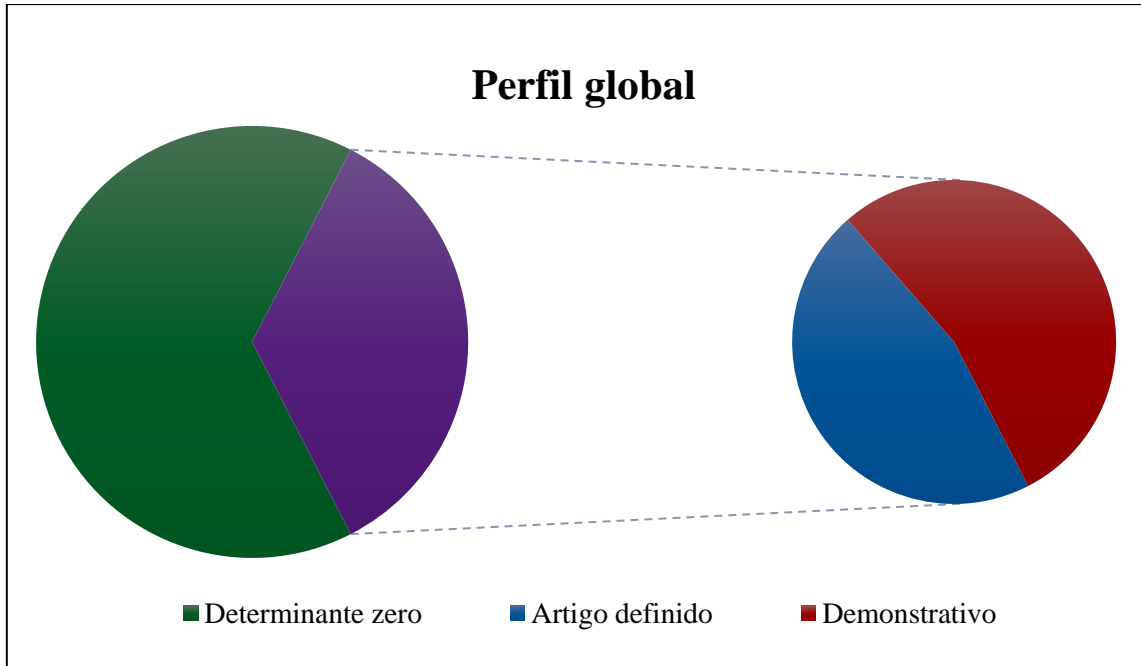
Em termos de números brutos, as ocorrências foram distribuídas em 520 SDs com determinante \emptyset , 65,1%, 150 com demonstrativos, 18,8%, e 129 com artigos definidos, 16,1%. A distribuição geral entre as formas de D pode ser melhor visualizada no Quadro 5:

Quadro 4 - Frequência das variantes da variável dependente na fala da comunidade de Monte Café

Variável dependente	Ocorrências/total	Frequência
Determinante \emptyset	520/799	65,1%
Demonstrativo	150/799	18,8%
Artigo definido	129/799	16,1%

Como explicado no capítulo 3, a análise a seguir irá tratar esta variação como realizações de uma variável dependente binária, que compreende o **determinante foneticamente realizado**, representado pelo <artigo definido> e o <demonstrativo>, e o **determinante zero**. Essa configuração pode ser vista no gráfico 1, e os resultados obtidos serão apresentados na seção 4.2.

Gráfico 1 - Perfil global da variável dependente na fala da comunidade de Monte Café



O arquivo <cell> gerado com todas as variáveis independentes inicialmente propostas revelou que algumas variáveis independentes apresentavam nocautes ou eram problemáticas por conta da insuficiência de dados ou da distribuição muito desigual dos dados, conforme explicitam Guy & Zilles (2007). Após a análise da distribuição geral, foram eliminados da análise os dados dos falantes MCMCGH3, MCCEQM3 e MCINRM3 (um homem e duas mulheres da faixa 3), em cujas entrevistas não foi realizado o artigo definido. Foram excluídos, ainda, os dados de MCMFNM2, mulher da faixa 2, que não usou o demonstrativo durante a entrevista. Além desses ajustes, foi necessária a exclusão da variável **nomes contáveis/não-contáveis** por manifestar no nível 1 do *step up* um valor estatisticamente não significativo⁵⁸, e pela distribuição desequilibrada dos dados⁵⁹. Por fim, dentro dos outros grupos de fatores, para melhorar a qualidade da análise estatística, optou-se por descartar as variantes com menos de 15 ocorrências, a saber:

⁵⁸ Run # 6, 2 cells: Convergence at Iteration 4; Input 0.651; Group # 5 -- c: 0.492, n: 0.572; Log likelihood = -516.129; Significance = 0.215.

⁵⁹ 90,4% de nomes contáveis e 9,6% de nomes não-contáveis.

- (i) No grupo **posição do SD na oração**, descartamos os fatores **voz passiva**⁶⁰ e **existencial**⁶¹, visto que somente 0,6% dos dados correspondem ao uso da voz passiva (6 ocorrências). O uso existencial, por sua vez, registrou apenas 12 dados, 1,5%.
- (ii) No grupo **familiaridade**, a variante **conhecimento geral**⁶² apresentou 9 dados (1,1% do total), sendo 7 com determinante zero e 2 com artigo definido.
- (iii) No grupo **número no SD**, a variante **leitura com possibilidade de plural**⁶³ teve 10 ocorrências (1,1%), sendo 8 com determinante zero, 1 com artigo definido e 1 com demonstrativo.
- (iv) No grupo outro **constituente no SD**, foram descartadas as variantes **advérbio locativo**⁶⁴, com 7 dados (0,8%): 1 com determinante zero, 1 com artigo definido e 5 com demonstrativo; **adjetivo pré-nominal**⁶⁵, com apenas 1 dado (0,1%), sem determinante foneticamente realizado; **adjetivo pós-nominal**⁶⁶, com 15 dados, dos quais 12 aparecem com determinante zero e 3 com artigo definido.

Na seção 4.2, serão apresentadas a distribuição geral da variável binária e os grupos selecionados pelo GOLDVARB-X por exercerem uma influência estatisticamente significativa sobre o uso do determinante zero.

4.2 PERFIL GLOBAL DA VARIÁVEL BINÁRIA E RESULTADOS DAS ANÁLISES: GRUPOS SELECIONADOS PELO GOLDVARB-X.

Após a revisão e consolidação, a base de dados registrou 799 dados da variável dependente, distribuídos de forma binária: **determinante zero**, com 520 dados, percentual de 65% e **determinante foneticamente realizado**, com 279 dados, percentual de 35%. Essa distribuição pode ser melhor visualizada no Quadro 6 e no Gráfico 2.

⁶⁰ Exemplo: Quê dizê *a técnica* é dada memo por mim. [MCFACH1]

⁶¹ Exemplo: Então houve *paralisação* de trabalho. [MCAJFH2]

⁶² Exemplo: ...como nós companhamos *a Biblia*, em tempo padre já ñ contava assim... [MCAJFH2]

⁶³ Exemplo: Quando eu comecei logo logo a praticar, só um ferrão *a mão* inflamava-se muito... [MCFACH1]

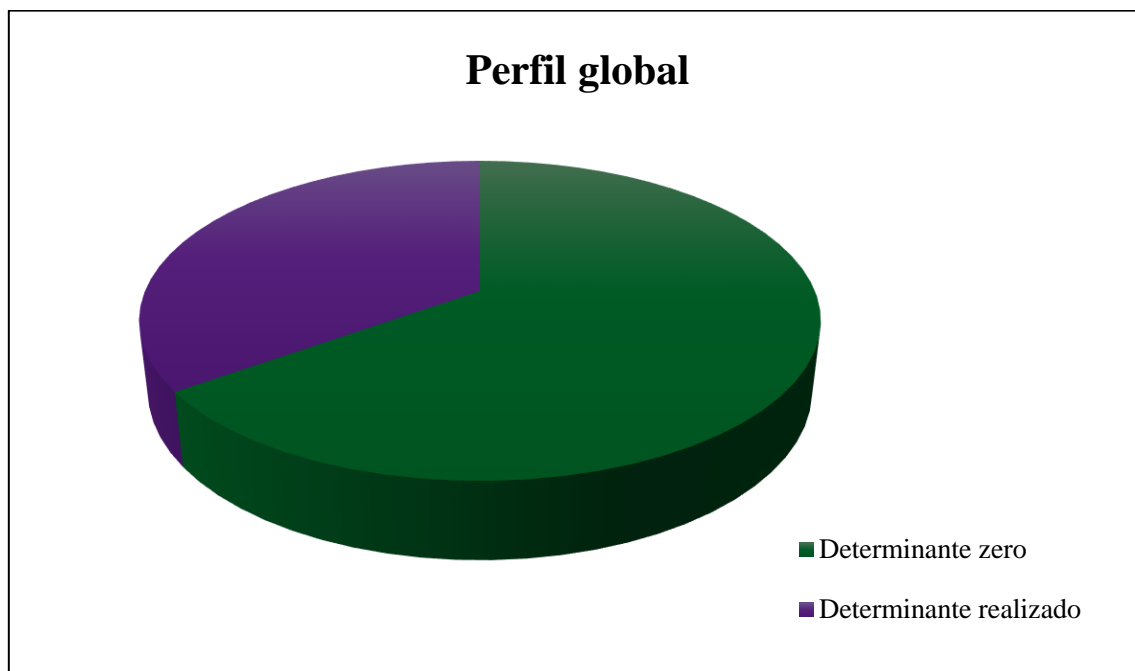
⁶⁴ Exemplo: Porque... segundo... disseram que... fez *essa aldeia aqui*.. só p'ra tonga. [MCLVLM1]

⁶⁵ Exemplo: Quê zê todos tem *mesma cor*, aqui São Tomé só xisti um tipo de macaco, só uma cor. [MCPFIH3]

⁶⁶ Exemplo: Talvez pessoa prope tem um dia que *prossor prope* diz, bom como eu sou tonga assim chega. [MCJCAH1]

Quadro 5 - Frequência da variável dependente na fala da comunidade de Monte Café

Variável dependente	Ocorrências/total	Frequência
Determinante Ø	520/799	65,1%
Determinante foneticamente realizado	279/799	34,9%

Gráfico 2 - Perfil global da frequência da variável dependente binária na fala da comunidade de Monte Café

A fim de analisar a realização e o tipo de determinante [+ específicos][+definidos] nos sujeitos e objetos diretos na comunidade de Monte Café, foram avaliados os efeitos de 9 variáveis independentes: 7 linguísticas e 2 extralinguísticas. O Quadro 7 apresenta as variáveis selecionadas pelo GOLDVARB-X como fatores condicionantes da realização do determinante zero.

Quadro 6 - Variáveis independentes selecionadas e rejeitadas pelo GOLDVARB-X

Variáveis independentes selecionadas	Ordem de seleção
Número no SD	+1
Presença de outro constituinte	+2
Faixa etária	+3
Sexo	+4

Familiaridade	+5
Animacidade do substantivo	+6
Variáveis independentes rejeitadas	Ordem de rejeição
Função sintática do SD	-1
Tipo de posse	-2
Posição do SD na oração	-3

A próxima seção apresenta os pormenores dos grupos selecionados, seguindo a ordem de seleção.

4.2 GRUPOS SELECIONADOS

4.2.1. Número do SD

Esta variável está estruturada em três fatores:

(i) Singular.

- (1) É, que **mãe** também veio d'Angola. [MCFJBH1]
- (2) É isso aqui lapaz fê jadinelo, isso lapaz isso **isso criança** aqui é mas bô. [MCALEH2]

(ii) Plural.

- (3) Sempre, ela me dava essa coragem, **cunhadas** também me dava essa coragem sempre... eu tô... até hoje. [MCLVLM1]
- (4) É tavam lá ele vivia aqui com tio, **os pai** vieram cá passá natal, quando foram levaram. [MCJCAH1]

(iii) Singular com leitura de plural.

- (5) Nó dava conta das criança como é devido... **Criança** comia mal... fazia tôda essa coêsa... [MCJNKM1]
- (6) Então lava ele **o pé**, caza, essas coisa assim... [MCJNKM1]

A Tabela 1 e os Gráficos 3 e 4 apresentam os resultados da análise do efeito do **número no SD**:

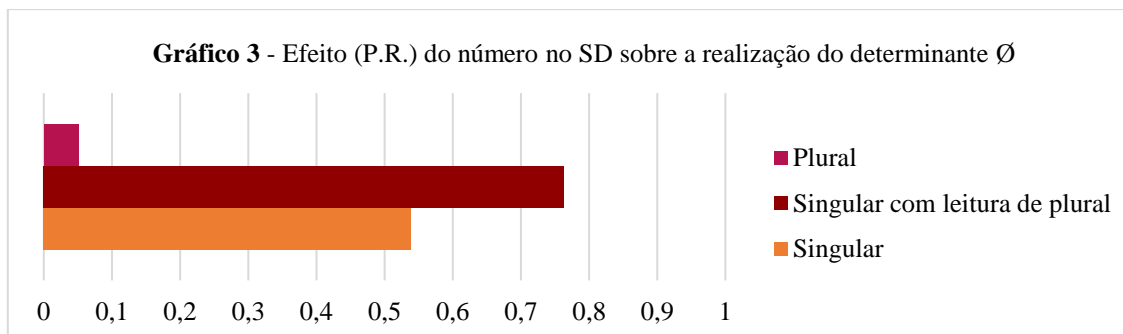
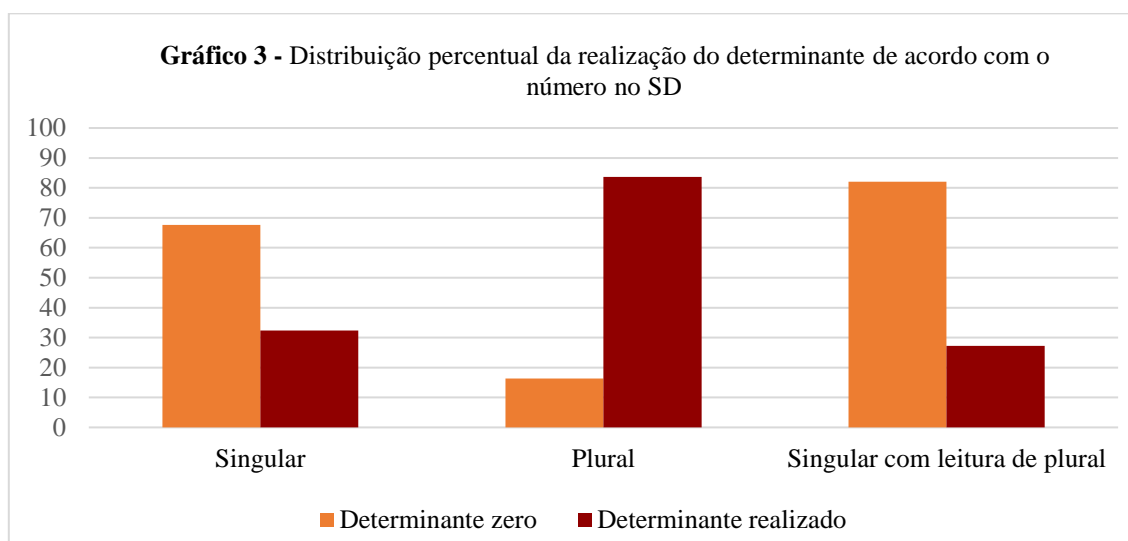


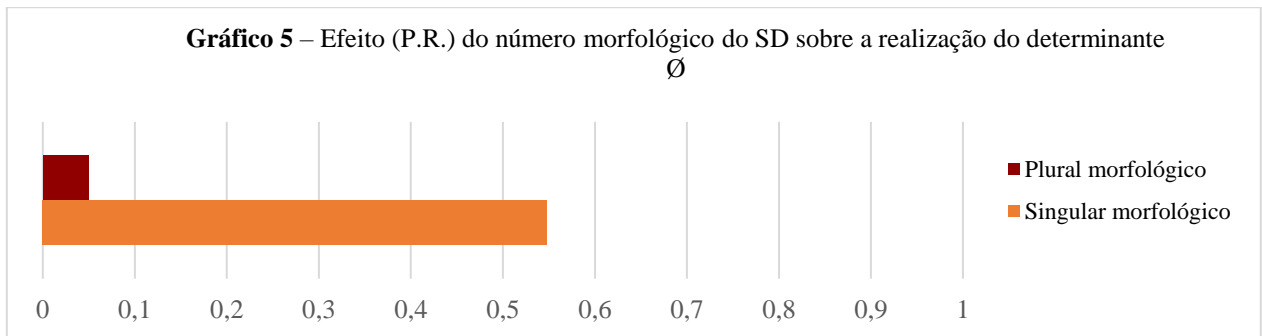
Tabela 1 - Efeito do número no SD sobre a realização do determinante Ø

Número no SD	Determinante Ø	Determinante realizado	Peso relativo	Total
Singular	483/714 67,6%	231/714 32,4%	0.538	714/791 90,3%
Plural	8/49 16,3%	41/49 83,7%	0.052	49/791 6,2%
Singular com leitura de plural	23/28 82,1%	5/28 17,9%	0.771	28/791 3,5%
Total	514/791 65%	277/791 35%		

Run # 39, 169 cells; Convergence at Iteration 11; Input 0.684; Log likelihood = -407.222; Significance = 0.017



Os resultados revelam que quando o SD é **singular** favorece o uso do determinante \emptyset , com peso relativo de 0.538 e frequência de 90%. Quando o SD é **singular com leitura de plural**, o determinante \emptyset é ainda mais favorecido, com peso relativo de 0.771 e frequência de 82%. Contudo, há apenas 28 dados nesse grupo, que correspondem a meros 3,5% do total do *corpus*. Quando o SD tem morfologia de **plural**, por sua vez, o uso do determinante \emptyset é fortemente desfavorecido, com peso relativo de 0.052 e frequência de 16%. Dessa maneira, confirma-se a hipótese de que os SDs **singulares**, mesmo os que possuem semântica de plural, favorecem o determinante \emptyset , enquanto o **plural** o desfavorece. Esse resultado é bastante esperado, visto que a marcação de **plural** no PMC, como no PB em geral, ocorre preferencialmente no determinante. A preferência pelo determinante fonético reflete, assim, essa possibilidade de marcação de plural no DET. Por outro lado, o favorecimento do determinante \emptyset nos SDs **singulares com leitura de plural** corrobora o constatado no PA de que a semântica plural parece não ser suficiente para favorecer a realização fonética do determinante (FIGUEIREDO, 2019, p. 376-377). A análise dos dados do PABH feita por Neves (2019, p. 184-186) mostra que, nos ODs [+ específicos][+ definidos], os SDs **singulares** favorecem o determinante \emptyset , contrastando com a variável **plural + leitura de plural**, que o desfavorece. A junção dos fatores **plural** e **leitura de plural** em Neves (2019), porém, não nos permite visualizar o efeito dos fatores semântico e sintático separadamente na realização desse determinante. Optamos, assim, por contrastar os usos no **singular morfológico** (singular, singular com leitura de plural e singular com possibilidade de plural⁶⁷) e no **plural morfológico**, a fim de avaliar o efeito do **número morfológico no SD** se comporta sobre a variável dependente. Os resultados estão na Tabela 2 e nos Gráficos 5 e 6:

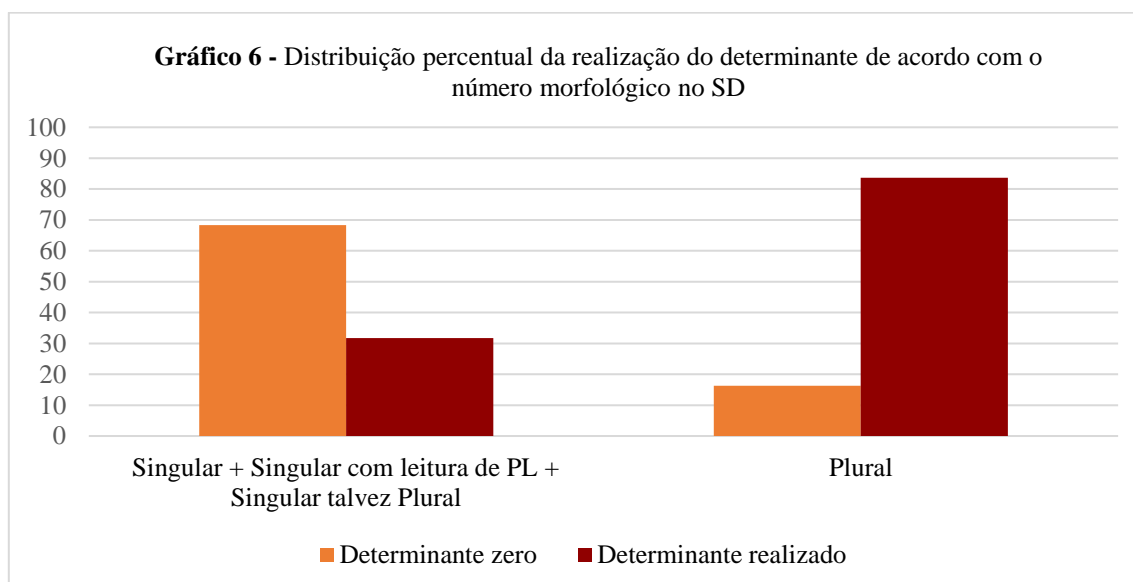


⁶⁷ Esse fator tinha sido retirado da análise por conter menos de 15 ocorrências.

Tabela 2 - Efeito do número morfológico no SD sobre a realização do determinante Ø

Número no SD	Determinante Ø	Determinante realizado	Peso relativo	Total
Singular + Singular com leitura de plural + Singular talvez plural	512/750 68,3%	238/750 31,7%	0.548	750/791 93,9%
Plural	8/49 16,3%	41/49 83,7%	0.050	49/791 6,1%
Total	520/799 65,1%	279/799 34,9%		

Run # 39, 152 cells; Convergence at Iteration 10; Input 0.684; Log likelihood = -409.326; Significance = 0.007



Esta segunda análise, com foco na morfologia e apenas dois fatores na variável independente, apresenta um resultado que agrega à análise anterior, voltada ao aspecto semântico: o uso do SD **singular** favorece levemente o determinante Ø, com peso relativo de 0.548 e frequência de 93,9%, ao passo que o SD **plural** o desfavorece fortemente, com peso relativo 0.050 e frequência de 6,1%. Contudo, o nível de significância do conjunto de fatores selecionados, 0.007, é melhor do que na análise inicial.

Além de encontrar paralelos nas referidas pesquisas sobre variedades de português, os resultados registrados na análise atual encontram semelhanças em estudos sobre o uso dos determinantes no SD do CCV e do ST. No CCV, a pluralidade é geralmente marcada através de

quantificadores, demonstrativos, possessivos ou numerais, enquanto os SDs singulares definidos são realizados como SDs nus (BAPTISTA, 2007, p. 82-84). Sobre os SDs singulares nus com semântica de plural, Lucchesi (1993)⁶⁸ afirma que tendem a ter interpretação [+ específica][+ definida] quando ocupam a função de sujeito; quando são ODs, costumam ser [- definidos][- específicos]. No ST também há uma relação direta entre definitude e pluralidade: os dois traços são atribuídos ao SD, ao mesmo tempo, através do determinante definido *inen*; os singulares definidos, em contrapartida, são normalmente SDs nus (ALEXANDRE; HAGEMEIJER, 2007, p.40-41).

4.2.2 Presença de outro constituinte no SD

Essa segunda variável selecionada está dividida em 5 variantes:

(i) Nenhum outro modificador.

(7) **Coração** trabalhava mal. Isso... sempre comprimidos aqui... [MCLVLM1]

(8) Um inviduo també no meio atiravam dum lado pa outro enquanto outro ia dis... disviando a bola e colocando **a lata**. [MCJCAH1]

(ii) Possessivo pós-nominal.

(9) **Mãe dela** que veio, quê a mia avô. [MCJCDH2]

(10) Vieram, **os pais dele** é que vieram. [MCJCAH1]

(iii) Sintagma preposicionado.

(11) **Copa do mundo** normalmente a gen vê, gente torce por país que(a) que fala nossa língua. [MCJCAH1]

(12) Só pra tirare isso, uma ve, é aquele uhora memo gente corta **isso pedaço di carne**... [MCSGPM3]

(iv) Oração relativa.

(13) **Adimistradoro que eu contrô** nã dexava sabê né livro né pegâ no livro. [MCPFIH3]

⁶⁸ apud Baptista (2007, p. 76).

(14) Agora **isso uhome quele casô** pariu cu ele fio. Tá lá. Ele tá lá cu fio dere tudo, tá ni'Sboa.

[MCCEQM3]

(v) Possessivo pré-nominal.

(15) Não consigui concluir porque **meu avô** não tinha condições. [MCJCAH1]

(16) **A minha mãe** ponha-me no colo ia pra o paiol. [MCFACH1]

A Tabela 3 e os Gráficos 7 e 8 apresentam os resultados do efeito da variável **presença de outro constituinte no SD** na realização do determinante zero.

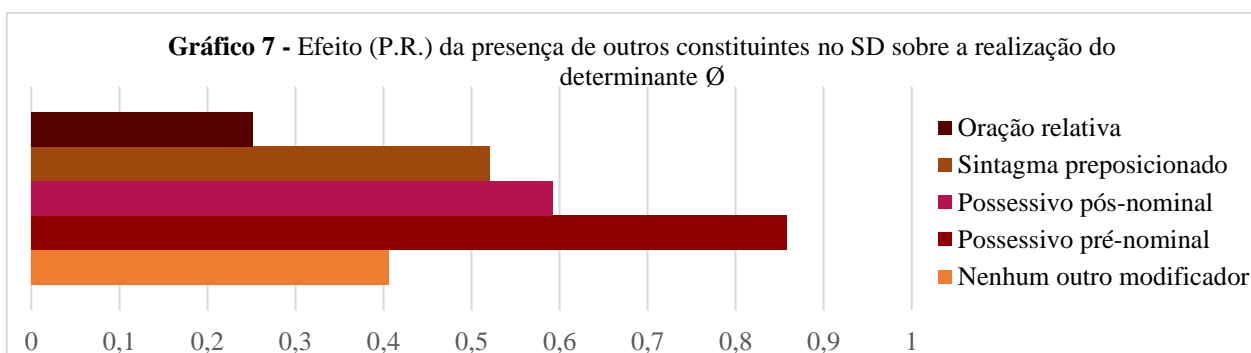
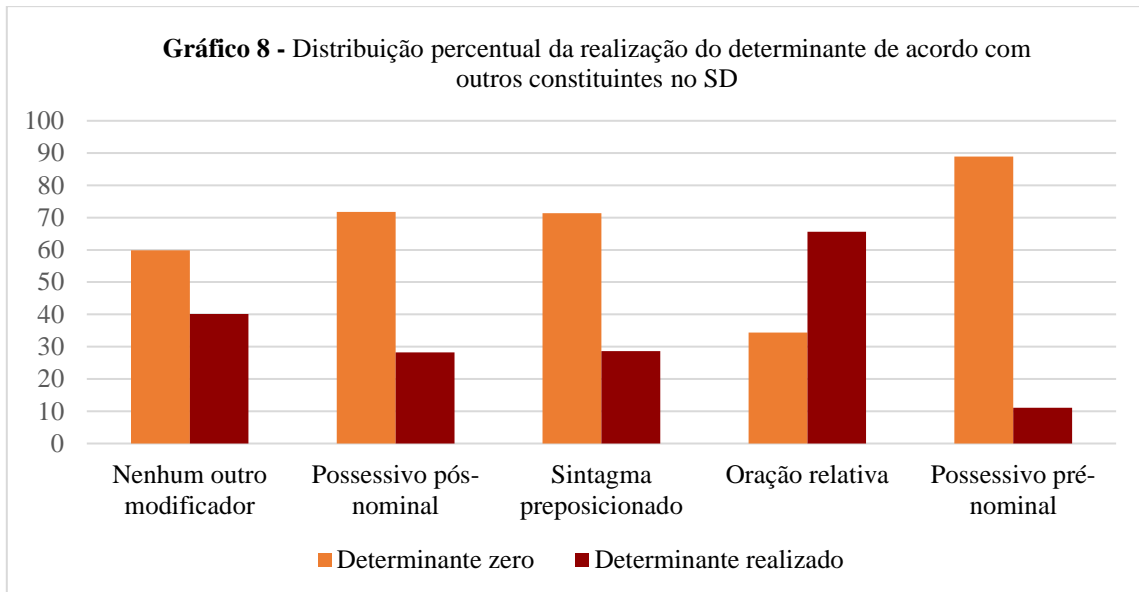


Tabela 3 - Efeito da presença de outros constituintes no SD sobre a realização do determinante \emptyset

Outro constituinte no SD	Determinante \emptyset	Determinante realizado	Peso relativo	Total
Nenhum outro modificador	306/511 59,9%	205/511 40,1%	0.403	511/755 69,1%
Possessivo pós-nominal	28/39 71,8%	11/39 28,2%	0.612	39/755 5,3%
Sintagma preposicionado	40/56 71,4%	16/56 28,6%	0.535	56/755 7,6%
Oração relativa	11/32 34,4%	21/32 65,6%	0.237	32/755 4,2%
Possessivo pré-nominal	104/117 88,9%	13/117 11,1%	0.860	117/755 15,8%
Total	489/755 64,8%	266/755 35,2%		



Dentre os constituintes presentes no SD, três favorecem o determinante \emptyset . O **possessivo pré-nominal** é o mais favorável, com peso relativo 0.860 e frequência de 89%. Em seguida, a presença de um **possessivo pós-nominal** exerce um favorecimento moderado, com peso relativo 0.612 e frequência de 71,8%. A presença de um **sintagma preposicionado** favorece o determinante \emptyset com peso relativo leve de 0.535 e frequência de 71,4%. Por outro lado, a **ausência de outros constituintes** no SD desfavorece levemente o determinante zero, com peso relativo 0.403 e frequência de 60%, enquanto a **oração relativa** inibe fortemente o uso do determinante zero, com peso relativo de 0.237 e frequência de 34%. Esses resultados apontam, em geral, o efeito inibidor que a presença de outros elementos atribuidores de referência possui sobre o artigo definido, constatado por Lucchesi (1993)⁶⁹ no CCV. A presença dos **possessivos pré-nominais** e dos **sintagmas preposicionados**, constituintes que favorecem o determinante \emptyset no PMC, se comporta de forma semelhante no PABH (BAXTER; LOPES, 2009, p. 326; NEVES, 2019, p. 175-176). Baxter e Lopes (2004)⁷⁰ observam, ainda, que os possessivos pré-nominais têm um comportamento muito específico, com um grau de referência definida suficiente para inibir a realização do artigo definido no PB. Os **possessivos pós-nominais** não são tratados nos estudos mencionados, mas também favorecem o determinante \emptyset no PMC. Por outro lado, o desfavorecimento do determinante

⁶⁹ apud Baxter e Lopes (2009, p.326).

⁷⁰ apud Baxter e Lopes (2009, p.326).

zero nos SDs com **orações relativas**, bem como nos **SDs sem outros constituintes**, também foi encontrado no PABH (BAXTER; LOPES, 2009, p. 326; NEVES, 2019, p. 176). No PA, a **presença de outros constituintes no SD** não foi considerada estatisticamente significativa como fator condicionante da variação do uso dos determinantes (FIGUEIREDO, 2019, p. 365).

4.2.3 Faixa etária

A faixa etária, terceira variável selecionada, permite rastrear o desenvolvimento do sintagma determinante em posição de sujeito e objeto [+ específico] [+ definido] no PMC através do tempo aparente, a fim de investigar se há, na comunidade, um perfil de aquisição do artigo definido, indicando uma mudança em curso. Entre os informantes de MC, esse fator está intrinsecamente ligado à escolaridade, visto que os falantes da faixa 3, em geral, não tiveram acesso ao ensino formal; alguns falantes da faixa 2, por sua vez, afirmam ter estudado por alguns anos; e todos os falantes da faixa 1 foram, em algum momento, à escola. A Tabela 4 e os Gráficos 9 e 10 apresentam os resultados do efeito da variável **faixa etária** na realização do determinante zero.

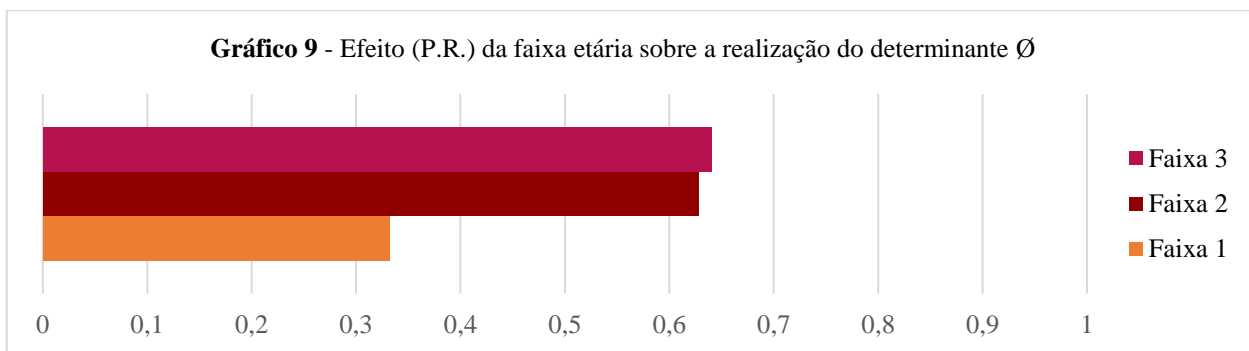
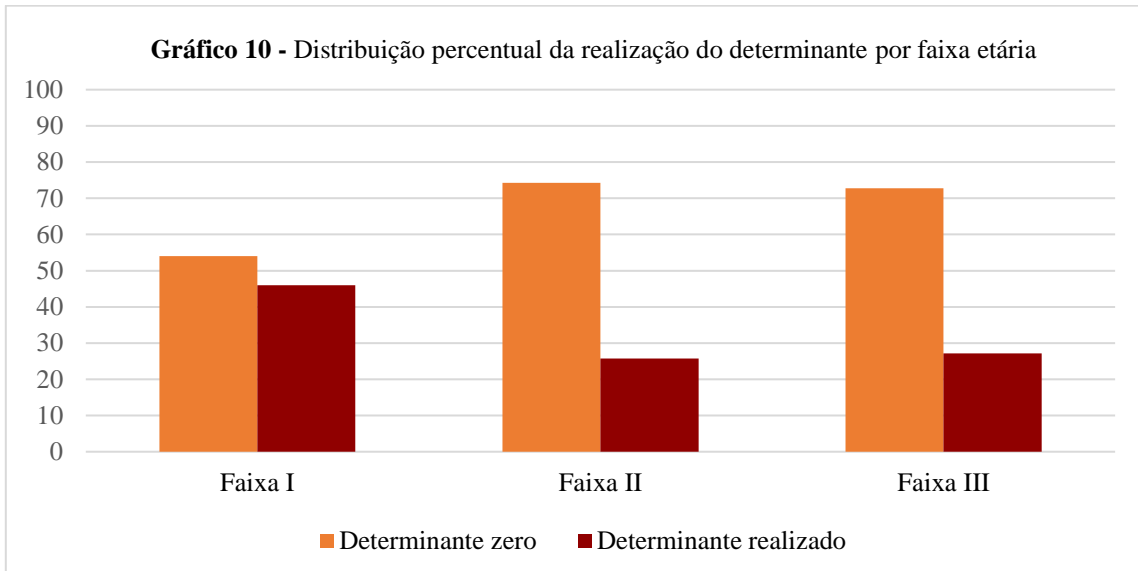


Tabela 4 - Efeito da faixa etária sobre a realização do determinante Ø

Faixa etária	Determinante Ø	Determinante realizado	Peso relativo	Total
Faixa 1	189/350 54%	161/350 46%	0.327	350/799 43,8%
Faixa 2	205/276 74,3%	71/276 25,7%	0.630	267/799 34,5%
Faixa 3	126/173 72,8%	47/173 27,2%	0.649	173/799 21,7%

Total	520/799 65,1%	279/799 34,9%
-------	------------------	------------------

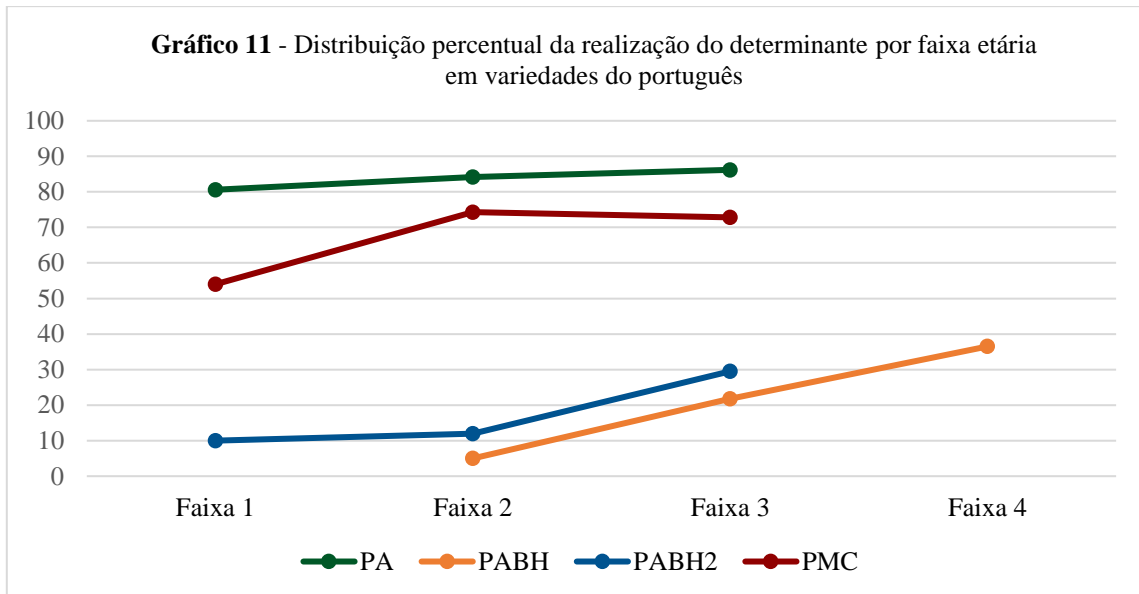
Run # 39, 169 cells; Convergence at Iteration 11; Input 0.684; Log likelihood = -407.222; Significance = 0.017



Observa-se que, na faixa 2 (41 a 60 anos) e na faixa 3 (> 60 anos), o determinante \emptyset é favorecido, o que não ocorre na faixa 1. Na **faixa 3**, o peso relativo é de 0.649 e frequência de 72,8%. Na **faixa 2**, o peso relativo é de 0.630 e a frequência de 74,3%. Na **faixa 1**, em contrapartida, o determinante \emptyset é moderadamente desfavorecido, com peso relativo de 0.327 e frequência de 54%. Considerando os resultados em tempo aparente, pode-se perceber um potencial desenvolvimento do sistema de determinante definido na comunidade. À primeira vista, o perfil aquisicional não se mostra tão nitidamente quanto no PABH (BAXTER; LOPES, 2009, p. 322; NEVES, 2019, p. 171) e se assemelha ao PA (FIGUEIREDO, 2019, p. 372), em que também há um aparente nivelamento entre as faixas 2 e 3. A frequência do uso do determinante \emptyset no PMC e nos estudos mencionados⁷¹ pode ser visualizado no Gráfico 11⁷².

⁷¹ No PABH, a distribuição das faixas etárias é a seguinte: faixa 2 (35 a 45 anos), faixa 3 (60 a 70 anos) e faixa 4 (mais de 80 anos) (BAXTER; LOPES, 2009, p. 322). No PA (FIGUEIREDO, 2019, p. 372), no PABH2 (NEVES, 2019, p. 171-172) e no PMC, as faixas etárias têm a mesma distribuição: faixa 1 (20 a 20 anos), faixa 2 (41 a 60 anos) e faixa 3 (> 60 anos).

⁷² Baxter e Lopes (2009, p. 322) e Figueiredo (2019, p. 372) analisaram o efeito da faixa etária sobre a realização do artigo definido. Para a constituição desse quadro comparativo, foi usada como referência a frequência da realização do determinante zero, calculando-se a porcentagem da diferença entre o total de ocorrências e a realização do artigo definido.



Essa diferença se deve, provavelmente, à maior proximidade geracional entre os falantes dos *corpora* de São Tomé e o período de transmissão irregular, em que havia fortes processos de contato linguístico e presença constante do português-L2. Vale salientar, ainda, que três entrevistas da faixa 3 foram retiradas da análise porque os falantes não realizaram o artigo definido – apenas o demonstrativo e o determinante zero. Em processos de aquisição em situação de contato linguístico, o demonstrativo, por ser foneticamente mais saliente, é adquirido antes do artigo definido.⁷³

4.2.4 Sexo

O sexo, quarta variável selecionada, é importante para se compreender a interferência da distribuição dos papéis sociais na língua falada na comunidade. Em MC, o acesso à comunidade externa é restrito para ambos os sexos, mas “os serviços dos homens os colocam em maior contato com os colonos e, conseqüentemente, em maior exposição à variante padrão” (SANTANA, 2019, p. 87). A Tabela 5 e os Gráficos 12 e 13 apresentam os resultados do efeito da variável **sexo** na realização do determinante zero.

⁷³ Tanto que nenhuma língua crioula de base portuguesa incorporou os artigos definidos do português, mas todas utilizam o demonstrativo para referência [+ definida][+ específica] (HASPELMATH, 2013).

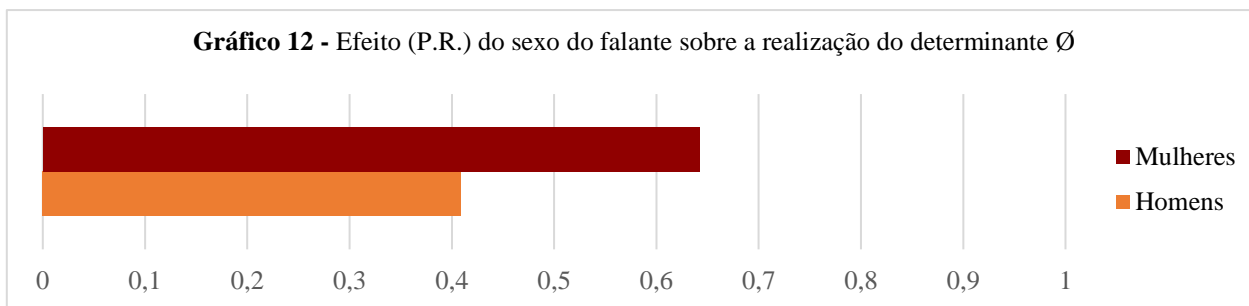
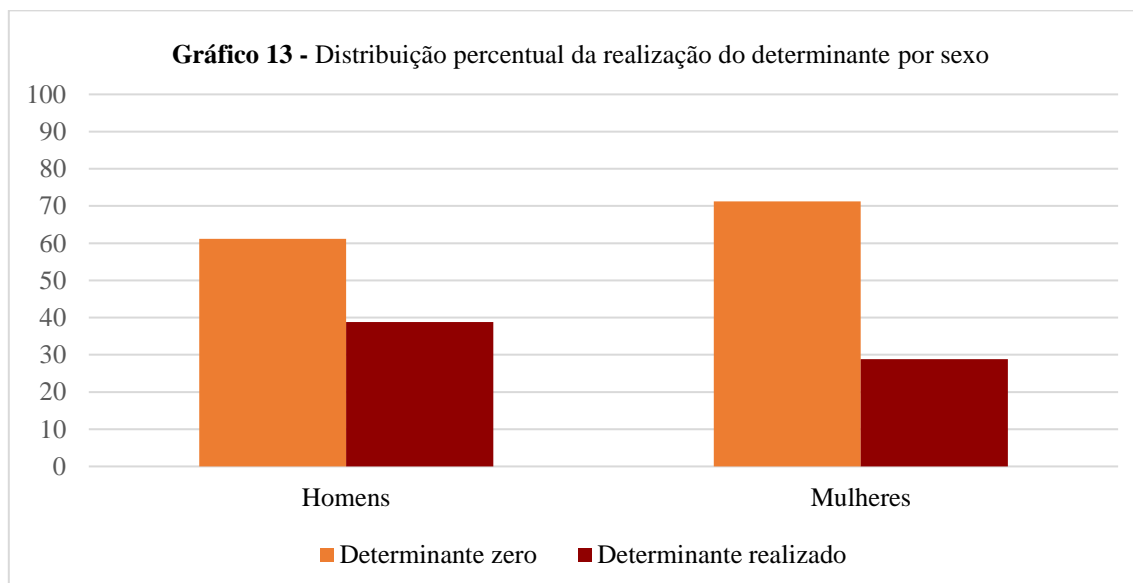


Tabela 5 - Efeito do sexo sobre a realização do determinante Ø

Sexo	Determinante Ø	Determinante realizado	Peso relativo	Total
Homens	300/490 61,2%	190/490 38,8%	0.411	490/799 61,3%
Mulheres	220/309 71,2%	89/309 28,8%	0.639	309/799 38,7%
Total	520/799 65,1%	279/799 34,9%		

Run # 39, 169 cells; Convergence at Iteration 11; Input 0.684; Log likelihood = -407.222; Significance = 0.017



O comportamento da variável sexo indica que na fala dos **homens** há um desfavorecimento do determinante Ø, com peso relativo 0.411 e frequência de 61,2%, enquanto a fala das **mulheres** o favorece, com peso relativo 0.639, com frequência de 71,2%. O resultado corresponde à hipótese

formulada: os homens utilizam mais a forma padrão em relação às mulheres, assim como acontece na variação dos Objetos Indiretos na mesma comunidade (SANTANA, 2019, p. 88). Na análise da variação dos determinantes do PA, essa variável se comportou de forma similar: as mulheres usam mais o determinante Ø e os homens utilizam mais a forma padrão (FIGUEIREDO, 2019, p. 371). No PABH, por sua vez, essa variável não é estatisticamente relevante (NEVES, 2019, p.189-194).

4.2.5 Familiaridade

A familiaridade, quinta variável selecionada, está distribuída em 4 variantes:

(i) Anafórica.

(17) Eu chamei capataz e **capataz** matou. [MCJNKM1]

(18) Então, aqui é bago... depois tem outra casca... gente quando secá tem que tirá **essa casca** ta qui em cima. [MCANJM1]

(ii) Referência cruzada.

(19) Homé també disanimô, disanimô co doença pega ni corda meté ni borso, fura daqui sartô jinela, chega lá embaxo caminho que va pa sede, sobe ponha corda ni cacau meté ni piscoço forró **cabeça**. [MCISMM2]

(20) Era um casal e também tinha só um casal de filho pra caso, foram muito bom. [...] **A senhora** trabalhava no coiso, trabalhava nos correio, o marido foi prussoro. [MCMFNM2]

(iii) Catafórica.

(21) **Dinhero** que gente ganha ã dá. Gen ganha muito pôco. [MCJCDH2]

(22) Eu sempre ã tive **esse hábito** quando trabalhava... matabichar depois trabalhar... [MCJNKM1]

(iv) Situacional.

(23) Venha pegare a mi aqui **graganta** tá gritare no pode. Yeu ficô mótira 'qui mia neto tudo tá chorare.... [MCCEQM3]

(24) Quando eu vim... eu encontrei **a senhora...** do seôro lá... em casa da mia irmã.
[MCJNKM1]

A Tabela 6 e os Gráficos 14 e 15 mostram os resultados do efeito da variável **familiaridade do SD** na realização do determinante zero quando [+ específico][+ definido].

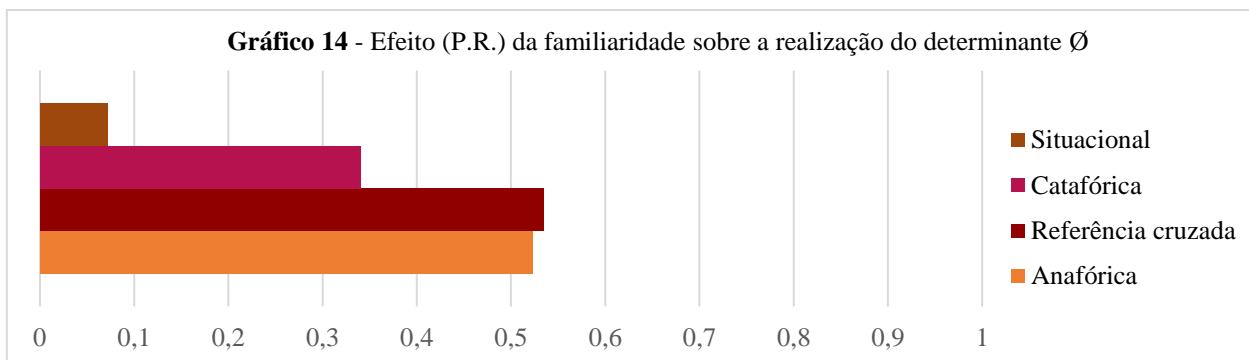
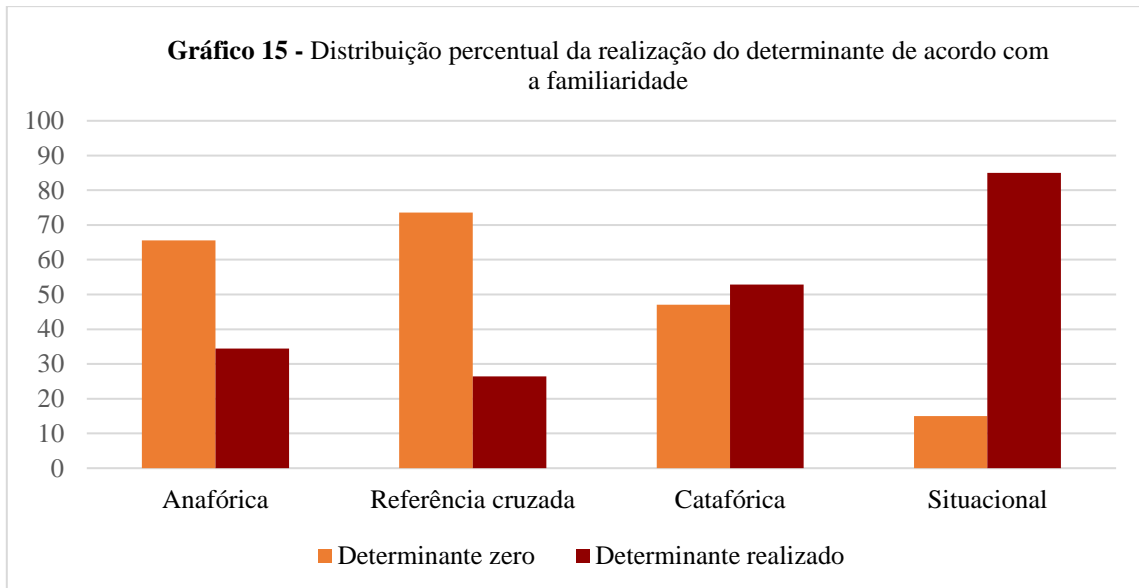


Tabela 6 - Efeito da familiaridade sobre a realização do determinante Ø

Familiaridade	Determinante Ø	Determinante realizado	Peso relativo	Total
Anafórica	391/596 65,6%	205/596 34,4%	0.524	596/790 75,4%
Referência cruzada	103/140 73,6%	37/140 26,4%	0.528	140/790 17,7%
Catafórica	16/34 47,1%	18/34 52,9%	0.344	34/790 4,3%
Situacional	3/20 15%	17/20 85%	0.075	20/790 2,5%
Total	513/790 64,9%	277/790 35,1%		

Run # 39, 169 cells; Convergence at Iteration 11; Input 0.684; Log likelihood = -407.222; Significance = 0.017



O comportamento dessa variável revela que quando a referência é anafórica ou cruzada, isso favorece levemente a realização do determinante \emptyset . No caso da **referência anafórica**, o peso relativo é de 0.524 e a frequência de 65,6%, enquanto as ocorrências com **referência cruzada** apresentam o peso relativo de 0.528 e a frequência de 73,6%. As referências catafórica e situacional, por sua vez, desfavorecem o uso do determinante zero. A referência **catafórica** apresentou o peso relativo de 0.344 e frequência de 47,1%. Por fim, a referência **situacional** mostrou um forte desfavorecimento, com peso relativo de 0.075 e frequência de 15%. Os resultados apontam a preferência pelo determinante \emptyset , também encontrada no PB e no CCV⁷⁴, quando o SD já foi mencionado na conversa (**anáfora**) (WALL, 2013, p. 237-238). O favorecimento no contexto de **referência cruzada** é comum ao PMC, ao PB⁷⁵ (WALL, 2013, p. 237-238) e ao PABH (NEVES, 2019, p. 180-183). No PABH, a familiaridade **situacional** também favorece a realização do determinante; por outro lado, no PABH a catáfora é a que mais favorece o determinante \emptyset (NEVES, 2019, p. 180-193), enquanto no PMC esse fator registra um desfavorecimento moderado.

⁷⁴ No CCV, Baptista (2007, p. 74-75) afirma que, quando a primeira menção é feita com marcação de definitude, a partir da segunda o seu uso é geralmente desnecessário.

⁷⁵ Wall (2013, p. 238) salienta que esses nomes, “embora não mencionados literalmente antes, são intimamente relacionados ao assunto do discurso e podem ser considerados como conceitualmente ‘presentes’ no discurso (por exemplo, *padre* ao falar sobre um serviço religioso, *princesa* ao fazer alusões a um conto de fadas, etc.)”. Tradução nossa.

4.2.6 Animacidade do substantivo

A animacidade do substantivo, sexta variável selecionada, está distribuída em duas variantes:

(i) Substantivo animado.

(25) **Pais dela** vieram, mãe também veio de lá pa São Tomé. [MCJCDH2]

(26) Nõ fui pra mato mais **os pais** é que nõ tinham possibilidade de me levar ao estudo. [MCFACH1]

(ii) Substantivo inanimado.

(27) Quê, os patrões é que enganavam-nos. Dividiu, criô **separação**. [MCAJFH2]

(28) Quando eu naceu isso fio sabe que tê **isso marca**. [MCALEH2]

A Tabela 7 e os Gráficos 16 e 17 mostram os resultados do efeito da variável **animacidade do substantivo** na realização do determinante Ø.

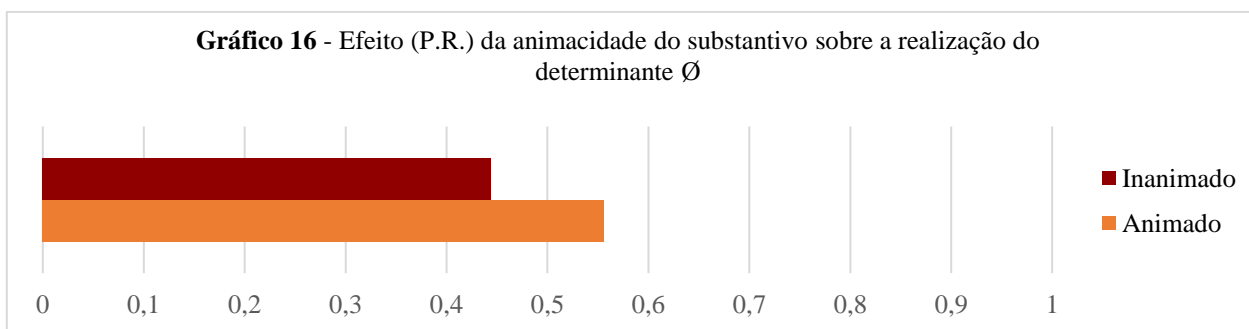
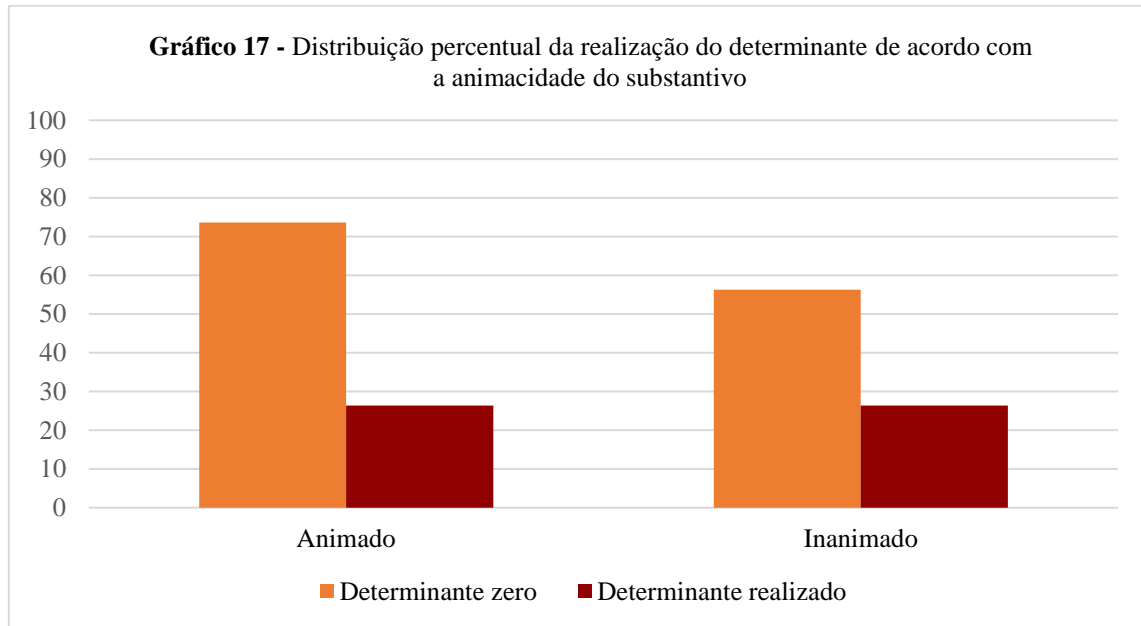


Tabela 7 - Efeito da animacidade do substantivo sobre a realização do determinante Ø

Animacidade do substantivo	Determinante Ø	Determinante realizado	Peso relativo	Total
Animado	293/398 73,6%	105/398 26,4%	0.576	398/796 50%
Inanimado	224/398 56,3%	174/398 43,7%	0.424	398/796 50%

Total	517/796 64,9%	279/799 35,1%
-------	------------------	------------------

Run # 39, 169 cells; Convergence at Iteration 11; Input 0.684; Log likelihood = -407.222; Significance = 0.017



Os resultados indicam que, quando o substantivo é **animado**, favorece moderadamente a realização do determinante \emptyset , com peso relativo de 0.576 e frequência de 73,6%. Por outro lado, os substantivos **inanimados** registram um peso relativo 0.424, com frequência de 56,3%, desfavorecendo, também moderadamente, o uso do determinante zero. O perfil dessa variável corresponde à proposta de Lyons (1999) de que o traço [+ humano], incluso no [+ animado], favorece o determinante \emptyset , resultado também encontrado no PA (FIGUEIREDO, 2019, p. 374). No PABH, a animacidade não é um fator considerado estatisticamente relevante no condicionamento do uso dos determinantes (BAXTER; LOPES, 2009, p. 322; NEVES, 2019, p. 194). No ST, relevante para o PA mas talvez minimamente para o PMC, o marcador de plural e definitude *inen* também mostra uma preferência por nomes [+ humanos], acompanhando-os sem a necessidade de outros especificadores, o que geralmente não acontece com os [- humanos], que requerem outros elementos marcadores que marquem especificidade (ALEXANDRE; HAGEMEIJER, 2007, p. 41).

4.3 GRUPOS NÃO SELECIONADOS PELO GOLDVARB-X

Nesta seção, apresentaremos os grupos rejeitados pelo GOLDVARB-X na análise estatística. Afinal, tanto a significância de um efeito quanto a falta dela são respostas às perguntas do pesquisador. Faz-se necessário, assim, mostrar não somente os resultados que validam as hipóteses inicialmente formuladas de condicionamento de uso, mas também os que contradizem essas pressuposições (GUY, ZILLES; 2007, p. 214). A apresentação dos grupos rejeitados pode (i) trazer progresso para as pesquisas que investigam a realização do determinante \emptyset , apontando que uma dada pergunta já foi feita e que a resposta para ela foi negativa; (ii) mostrar que há um efeito, ainda que leve, dos fatores envolvidos; (iii) revelar que os dados coletados são insuficientes para mostrarem a relevância de um dado efeito (GUY; ZILLES, 2007, p. 215). É papel do investigador ‘iluminar’ o fenômeno estudado o quanto puder, oferecendo o máximo de informações possível. A relevância dos grupos **função sintática do SD**, **tipo de posse** e **posição do SD na oração**, apresentados a seguir, se dá por acrescentar informações em relação ao que já se sabe a respeito da língua e do fenômeno enfocados nesta análise.

4.3.1. Função sintática do SD

A variável **função sintática do SD** está estruturada em duas variantes e foi a primeira a ser rejeitada pelo GOLDVARB-X, em todas as rodadas realizadas. Inclusive, quando retiramos das rodadas o grupo **posição do SD na oração**, devido a uma potencial redundância entre essas duas variáveis, ela continua sendo rejeitada.

(i) Sujeito.

(29) Dia que eu vô pa lá avó chora uma coêsa doida porque... **filho dela** moreu... porque nã sô quê... [MCANJM1]

(30) **A vida de sempre** sempre foi como é agora, dura, o que aparece é que gente usa. [MCJCAH1]

(ii) Objeto direto.

(31) Como lá é... eu sei... o zona de cobra... então não segurei saca não... chamei **capataz**.

[MCJNKM1]

(32) E e é muito normal, a evolução nós temo tomamos a **independência**, tamos na

democracia. [MCFACH1]

A Tabela 8 e o Gráfico 18 apresentam os pesos relativos registrados da variável **função sintática do SD** quando avaliada independentemente no nível 1 do *Stepup*. O Gráfico 19 apresenta a distribuição de frequências dos seus dois fatores.

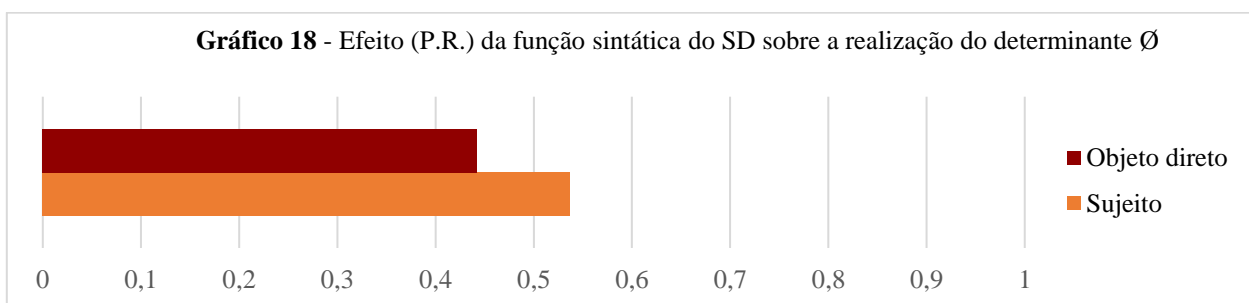
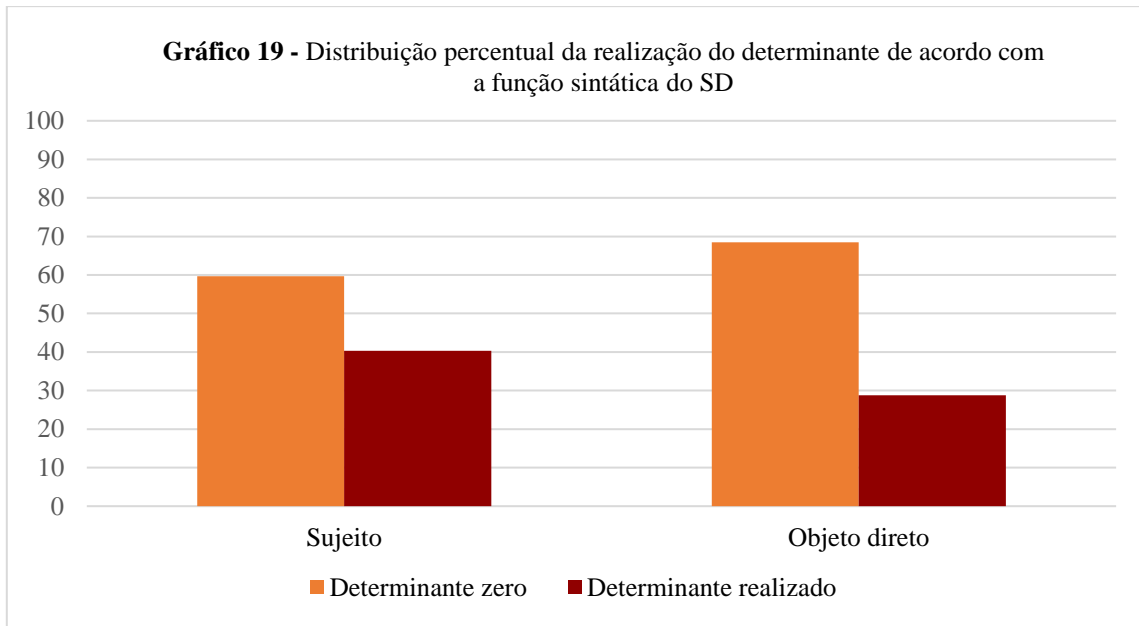


Tabela 8 - Efeito da função sintática do SD sobre a realização do determinante Ø

Função sintática do SD	Determinante Ø	Determinante realizado	Peso relativo	Total
Sujeito	333/486 68,5%	153/486 31,5%	0.537	486/799 60,8%
Objeto direto	187/313 59,7%	126/313 40,3%	0.442	313/799 39,2%
Total	520/799 65,1%	279/799 34,9%		

Level 1 stepup; Log likelihood = -513.702; Significance = 0.012



À primeira vista, há uma diferença sugestiva entre as frequências dos dois fatores. Os resultados no nível 1 sugerem que o **sujeito** favorece o determinante \emptyset , com peso relativo 0.537 e frequência de 68,5%, enquanto o **objeto direto** o desfavorece, com peso relativo 0.442 e frequência de 29,5%. No CCV, há um comportamento semelhante à aparente tendência da Tabela 8: os SDs nus singulares definidos tendem a aparecer em posição de **sujeito**. Baptista (2007, p. 77) observa que, quando na posição de sujeito, os SDs nus costumam trazer informações já conhecidas, são [+definidos], ao passo que, na posição de objeto, trazem informações novas e são, portanto, [-definidos]. No PABH, ao analisarem SDs exercendo 7 funções sintáticas⁷⁶, o **sujeito** e o **objeto direto** se comportaram de forma semelhante, favorecendo levemente o determinante \emptyset (BAXTER; LOPES, 2009, p. 328-329). Uma investigação mais apurada desse grupo em futuras análises, aumentando o *corpus* e refinando o conjunto de variáveis independentes (visto que apresentamos, para poucos dados, muitas variáveis independentes), pode manifestar a relevância da **função sintática do SD** na realização do determinante zero.

⁷⁶ Baxter e Lopes (2009, p. 329) investigaram o efeito das funções de *sujeito*, *objeto direto*, *predicativo*, *adjunto adverbial*, *adjunto adnominal*, *objeto indireto* e *frase independente* sobre a realização do artigo definido. Nesse estudo, todas as variantes, exceto o adjunto adverbial, favoreceram o SD nu.

4.3.2 Tipo de posse

A variável **tipo de posse**, a segunda rejeitada pelo programa de análise estatística, está distribuída em duas variantes:

(i) Posse alienável.

(33) 'Sinhoro, sô diretô!' Taqui **chave**, quando chega hora você fecha casa, chave vai nu isquitorio. [MCPFIH3]

(34) Gende di depedecha quando saio ni depedecha quando feze isso ardea, é que raçabê **isso casa**. Caté aqui. [MCINRM3]

(ii) Posse inalienável.

(35) ...porque **minha filha** é maior do [que o] primeiro... depois esse primeiro... teve ôtro filho... [MCLVLM1]

(36) Não **a família dele** vive aqui, lá na roça Piedade. [MCANJM1]

A Tabela 9 e o Gráfico 20 apresentam os pesos relativos registrados da variável **tipo de posse** quando avaliada independentemente no nível 1 do *Stepup*. O Gráfico 12 apresenta a distribuição de frequências dos seus dois fatores.

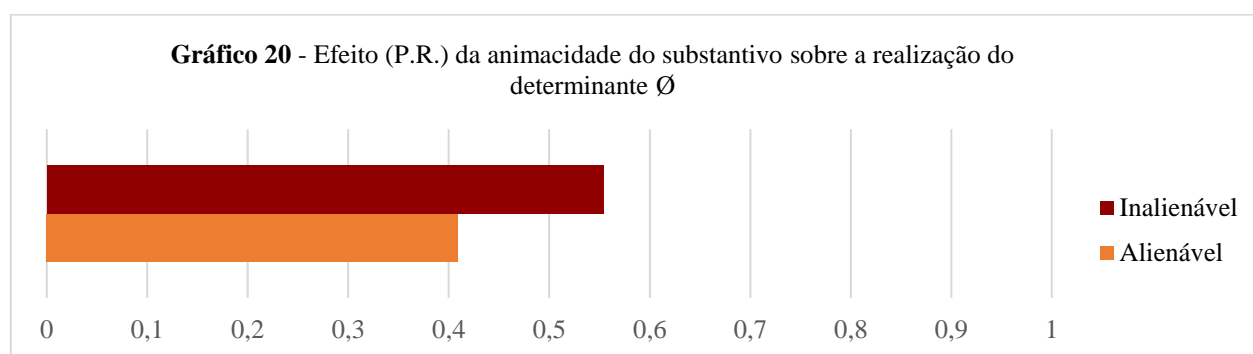
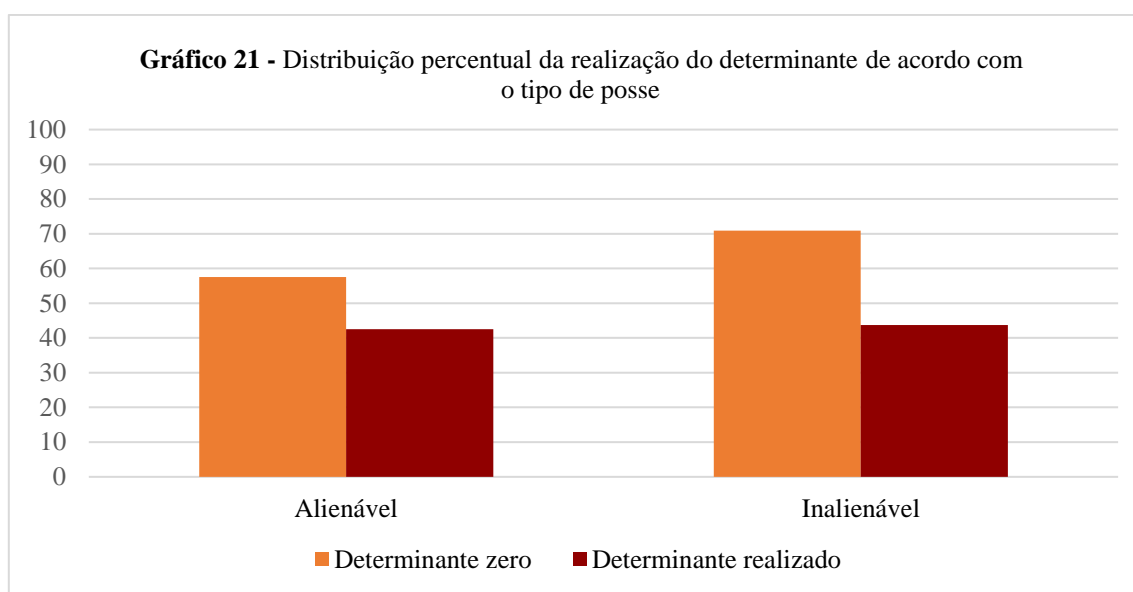


Tabela 9 - Efeito do tipo de posse sobre a realização do determinante \emptyset

Tipo de posse	Determinante \emptyset	Determinante realizado	Peso relativo	Total
Alienável	104/181 57,5%	77/181 42,5%	0.409	181/487 37,2%
Inalienável	217/306 70,9%	89/306 29,1%	0.554	306/487 62,8%
Total	321/487 65,9%	166/487 34,1%		

Level 1 stepup; Log likelihood = -512.299; Significance = 0.004



Neste caso também, as frequências e os pesos relativos registrados no nível 1 sugerem que esta variável independente poderia exercer um efeito. Nesse nível, a posse **inalienável** favorece o determinante \emptyset , com peso relativo 0.554, correspondendo a uma frequência de 70,9%, e a posse **alienável** o desfavorece, com peso relativo 0.409 e frequência de 57,5%. Embora pesos relativos sejam bastante modestos, estes resultados são consoantes com a hipótese baseada na proposta de Lyons (1999) de que a posse inalienável, por estar intrinsecamente ligada ao possuidor, favorece o SD nu. Tendências semelhantes foram encontradas no PABH, ainda que esse grupo de fatores também não seja considerado estatisticamente significativo (NEVES, 2019, p. 187-189). O perfil do **tipo de posse** sugere que essa variável deve ser reavaliada em futuras pesquisas, mais refinadas

e contemplando mais dados, a fim de verificar se esse grupo tem relevância na realização do determinante zero.

4.3.3 Posição do SD na oração

A última variável rejeitada pelo programa foi a **posição do SD na oração**, que está distribuída em 4 variantes. Essa variável foi rejeitada em todas as rodadas, mesmo quando o grupo **função sintática do SD** foi retirado.

(i) Sujeito em posição não-marcada.

(37) Bom, stado prutendeu de entra toma empresa, **stado** tomô empresa, stado pagô todo mês. [MCJCDH2]

(38) Cresci assim, **os tios** aqui tudo eram disso e levavam-nos domingos ocê era obrigado a ir. [MCFJBH1]

(ii) Sujeito em posição marcada.

(39) Trabalaiô aqui també cu ca catrapila aqui de de empresa, **rumão dele**. [MCMCGH3]

(40) tem uma filha que morreu afogado no no mar também **essa prussora**. [MCAJFH2]

(iii) Objeto direto.

(41) É. Tá sicuro. Abre **janela**. [MCCEQM3]

(42) Nõ vô consintir **a dor** pra nada, nõ consigo nõ consigo com consentir. [MCFACH1]

(iv) Foco.

(43) ...depois gente viu só a vir, só **cabeça** que gen viu. [MCAJFH2]

(44) Ele nõ sabe tocar violão, ele só pegô esse vio... essa viola... **esse violão** ele trouxe pra mi, ta comigo até hoje. [MCFJBH1]

A Tabela 10 e o Gráfico 22 apresentam os pesos relativos registrados da variável **posição do SD na oração** quando avaliada independentemente no nível 1 do *Stepup*. O Gráfico 13 apresenta a distribuição de frequências dos seus fatores.

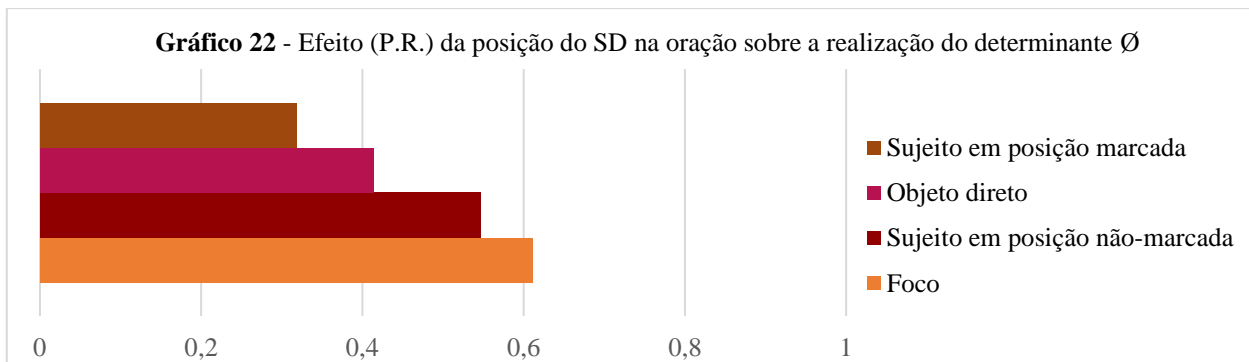
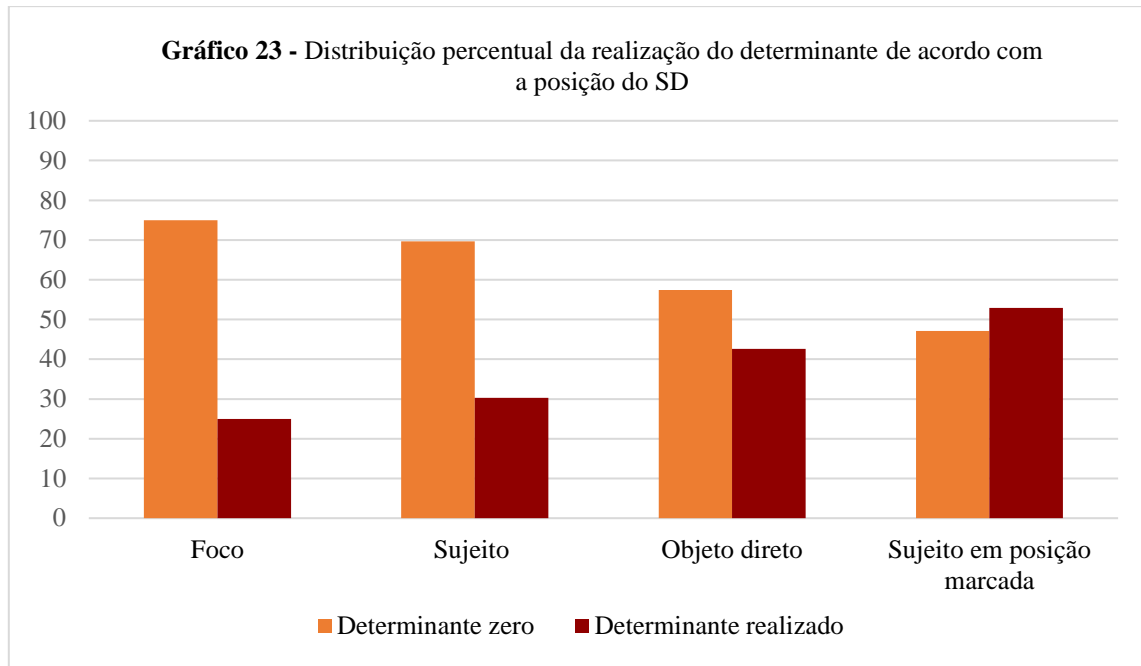


Tabela 10 - Efeito da posição do SD sobre a realização do determinante \emptyset

Posição do SD na oração	Determinante \emptyset	Determinante realizado	Peso relativo	Total
Foco	30/40 75%	10/40 25%	0.611	40/781 5,1%
Sujeito em posição não-marcada	320/459 69,7%	139/459 30,3%	0.547	459/781 58,8%
Objeto direto	152/265 57,4%	113/265 42,6%	0.414	265/781 33,9%
Sujeito em posição marcada	8/17 47,1%	9/17 52,9%	0.318	17/781 2,2%
Total	510/781 65,3%	271/781 34,7%		

Level 1 stepup; Log likelihood = -509.243; Significance = 0.003



O comportamento dessa variável mostra que, quando o SD antecede o verbo, o determinante \emptyset é favorecido, e quando o sucede, é desfavorecido. Na posição de **sujeito em posição não-marcada**, o peso relativo é de 0.547 e a frequência é de 69,7%, e quando se trata de um OD na posição de **foco**, o favorecimento é maior, com peso relativo 0.611 e frequência de 75%. Neste caso também, o perfil da variável independente testada independentemente no nível 1 do *Stepup* parece estar em consonância com a hipótese de que, como no PB, os SDs nus [+ definidos][+ específicos] ocorrem preferencialmente na posição pré-verbal, seja com a função sintática de sujeito ou de OD. O desfavorecimento do determinante \emptyset ocorre na posição de **objeto direto**, com peso relativo 0.414 e frequência de 57,4%, e de **sujeito em posição marcada**, com peso relativo de 0.318 e frequência de 47,1%. Esse grupo foi inicialmente inserido na análise dos OD do PABH, mas posteriormente retirado, devido à distribuição muito desigual entre as duas variantes (NEVES, 2019, p. 158). Espera-se que a relevância da **posição do SD na oração** para a realização do determinante zero pode ser melhor investigada em uma análise mais apurada, aumentando a quantidade de dados e diminuindo o desequilíbrio da sua distribuição entre das variantes.

4.4 SÍNTESE

Os resultados da análise do condicionamento revelam que os usos do determinante [+ específico][+ definido] em posição de sujeito e objeto direto é condicionado por fatores semânticos, morfossintáticos, pragmáticos e sociais. O favorecimento do Det Ø diante de outros constituintes no SD confirma que, na presença de elementos que marquem referencialidade, não é necessária a realização de um determinante fonético. A semântica singular, bem como sua estrutura morfossintática, também favorece a realização do Det Ø. A familiaridade da referência é significativa em orientar a forma da identificação do referente, que quando uma informação já é compartilhada com o interlocutor, prioriza-se o uso do Det Ø, diferentemente de quando a familiaridade com o constituinte se dá em seguida. Outro fator que se mostrou relevante foi a animacidade, confirmando a relação proposta por Lyons (1999) entre os traços [+ definido] e [+ animado] e o determinante nulo. Relativamente às variáveis independentes sociais propostas como potenciais condicionadoras do determinante zero, a faixa etária e o sexo se revelaram significativos, apontando, por um lado, para um comportamento mais conservador na fala das mulheres (com maior presença do Det Ø) em relação à dos homens, e, por outro lado, para um processo aquisicional em vigor, cuja ordem se dá pelos usos do determinante Ø > demonstrativo não-dêitico (foneticamente mais saliente) > artigo definido.

5. CONSIDERAÇÕES FINAIS

No decorrer desta dissertação, buscou-se investigar e descrever as condições semânticas, sintáticas e extralinguísticas que condicionam os SDs nus [+ específicos][+ definidos] em posição de sujeito e de objeto no Português de Monte Café. Os resultados obtidos na análise do *corpus* contemplado demonstram uma preferência pelo SD nu, com 65% de frequência, enquanto os outros 35% estão distribuídos de forma equivalente entre o demonstrativo e o determinante Ø. Na análise do comportamento da variável dependente binária (determinante zero e determinante foneticamente realizado), os grupos de fatores considerados estatisticamente relevantes para a realização do determinante Ø foram, por ordem de seleção: **número no SD, presença de outro constituinte no SD, faixa etária, sexo, familiaridade e animacidade do substantivo**.

Mostrou-se que os resultados confirmam, parcial ou completamente, as hipóteses propostas para os fatores selecionados. O efeito da variável **número no SD** confirmou que o singular favorece o SD nu, e as análises semântica e morfossintática apontam para resultados semelhantes. Por um lado, o SD singular morfológico inibe o determinante fonético, mesmo quando tem semântica plural. Por outro lado, a preferência pela realização desse determinante nos SDs morfológicamente plurais reflete a possibilidade de marcação de plural no determinante. A hipótese de que a **presença de outros constituintes no SD** condicionaria a realização do det Ø foi parcialmente confirmada. Em geral, a presença desses elementos inibe a realização do determinante fonético, mas a oração relativa a favorece. O estudo em tempo aparente, resultado do efeito da **faixa etária**, mostra um perfil aquicional, em que os falantes da faixa 2 e 3 utilizam preferencialmente o determinante Ø, e a faixa 1, mais escolarizada, favorece o determinante fonético. Isso parece indicar uma mudança no sistema de determinante na comunidade, em consonância com outros estudos sobre diferentes fenômenos linguísticos nesta comunidade (BAXTER, 2002, 2004; SANTANA, 2019) que constata grandes mudanças morfossintáticas em curso. Considerando que houve, em Monte Café, uma situação de transmissão linguística irregular (BAXTER, 1995; LUCCHESI; BAXTER, 2009), o determinante Ø e o demonstrativo não-dêitico são vestígios de fases iniciais da aquisição do português como L2, que foram incorporados na gramática L1 da comunidade. Nas fases iniciais da aquisição, no contexto de contato, teria havido muita variação a nível de língua-E, e essa variação serviu de modelo (dado linguístico primário) para a aquisição do português como L1 (BAXTER, 1995; LUCCHESI; BAXTER, 2009; LUCCHESI; RIBEIRO, 2009). Como a contratação de

trabalhadores continentais funcionou em ciclos de 5 anos desde meados do século XIX até pouco antes da revolução em Portugal em 1974, houve uma presença constante de falantes de português L2, alimentando constantemente o processo de aquisição de L1 e transmissão geracional. O efeito do **sexo** também confirmou que os homens, mais expostos à variedade padrão em contato com os colonos, são os que mais realizam o determinante fonético, enquanto a fala das mulheres favorece o determinante \emptyset . Finalmente, a hipótese de que a forma como o ouvinte identifica os SDs condiciona o uso do determinante zero foi confirmada. A **familiaridade** através da anáfora e da referência cruzada registrou um favorecimento do determinante \emptyset . A catáfora, bem como a presença do referente na situação, favorece o determinante fonético. Por fim, no grupo **animacidade**, confirmou-se que os substantivos animados, considerados mais salientes à percepção humana, favorecem o SD nu.

O estudo contribui, portanto, para a compreensão do desenvolvimento do SD em variedades atlânticas do português que passaram por fases de transmissão irregular, i.e., com dados linguísticos primários introduzidos a partir de uma fase de aquisição de português L2. Ademais, foram fornecidos dados que permitem uma caracterização mais detalhada do português tonga, junto a outros estudos que analisam outras variáveis morfossintáticas no português de Monte Café (BAXTER, 2002, 2004; SANTANA, 2019).

Por fim, são pertinentes algumas recomendações de futuros caminhos para pesquisas sobre o tema desta dissertação com dados da mesma comunidade. Nesse sentido, propõe-se: (i) uma análise quantitativa trinomial, a fim de observar o comportamento do artigo definido e do demonstrativo separadamente; (ii) o refinamento dos grupos de fatores e a exploração de cruzamentos de grupos, a fim de apurar a análise do comportamento do sujeito e do objeto; e (iii) um estudo em tempo real, a partir da constituição de um *corpus* com dados mais recentes da comunidade, que permita a comparação entre esses dois *corpora*.

REFERÊNCIAS

- ABNEY, S. P. *The English noun phrase in its sentential aspect*. Dissertação (PhD) – Massachusetts Institute of Technology, Massachusetts, 1987.
- ALEXANDRE, Nélia; HAGEMEIJER, Tjerk. Bare nouns and the nominal domain in Santome. In: BATISTA, Marlyse; GUERÓN, Jaqueline (Ed.). *Noun Phrase in Creole languages: a multi-faceted approach*. Amsterdam: John Benjamins, 2007. p.37-59.
- BAPTISTA, Marlyse. On the syntax and semantics of DP in Cape Verdean Creole. In: BAPTISTA, Marlyse; GUÉRON, Jacqueline (Ed.). *Noun phrases in creole languages: a multi-faceted approach*. Amsterdam; Philadelphia: John Benjamins, 2007. p.61-105.
- _____. *Noun phrases in creole languages. A multifaceted approach*. Amsterdam; Philadelphia: John Benjamins, 2007.
- BAXTER, Alan Norman. A contribuição das comunidades afro-brasileiras isoladas para o debate sobre a crioulação prévia: um exemplo do estado da Bahia. In: ANDRADE, E. (ed.). *Actas do Colóquio Internacional sobre Línguas Crioulas de Base Portuguesa*, Universidade de Lisboa. Lisboa: Colibri, p.7-35, 1992.
- _____. Transmissão Geracional Irregular na História do Português Brasileiro - divergências nas vertentes afro-brasileiras. *Revista Internacional de Língua Portuguesa*, n.14, p. 72-90, 1995.
- _____. *Semi-creolization: testing the hypothesis against data from Portuguese-derived languages of São Tomé (West Africa)*. Projeto financiado pelo Australian Research Council e sediado na La Trobe University, Melbourne, Austrália, 1998-2000.
- _____. Semicreolization? - The restructured Portuguese of the Tongas of São Tomé, a consequence of L1 acquisition in a special contact situation. *Journal of Portuguese Linguistics*, n.1, p. 7-39, 2002.
- _____. The development of variable NP plural agreement in a restructured African variety of Portuguese. In: Escure, Geneviève; Schwegler, Armin (Org.). *Creoles, Contact and Language change: Linguistics and social implications*. Amsterdam: John Benjamins Publishing Co, v.27, p. 97-126, 2004.
- BAXTER, Alan; LOPES, Norma. O artigo definido. In: LUCCHESI, D., BAXTER, A., and RIBEIRO, I. (Orgs). *O português afro-brasileiro*. Salvador: EDUFBA, p. 319-330, 2009.

- BAXTER, Alan; MELLO, Camila Ferreira de; SANTANA, Natali Gomes de Almeida. A construção de objeto duplo e as influências do substrato no português afrobrasileiro (e africano). *Revista PAPIA*, v. 24, n.2, p. 283-306, 2014.
- BEVILÁQUA, Kayron Campos. *Sintagmas nominais nus: um experimento sobre a distinção contável-massivo no português brasileiro*. Dissertação (Mestrado) – Universidade Federal do Paraná, Curitiba, 2015.
- CAMACHO, Roberto; PEZATTI, Erolilde. Nomes contáveis e não contáveis. In: *Alfa*. São Paulo, v.40, p. 59-74, 1996.
- CAVALCANTI Mariano, Victor. *Nomes nus na posição e sujeito no português afro-brasileiro*. Dissertação (Mestrado) – Universidade Federal da Bahia, Salvador, 2018.
- CHIERCHIA, Gennaro. Reference to Kinds across Languages. *Revista Natural Language Semantics*, n.6, p. 339-405, 1998.
- CHOMSKY, Noam. *Syntactic structures*. The Hague: Mouton, 1957.
- _____. Remarks on nominalization. In: JACOBS, Roderick A.; ROSEMBAUM, Peter S. (eds.) *Reading in English Transformational Grammar*. Massachusetts, p.184-221, 1970.
- FIGUEIREDO, Carlos Filipe Guimarães. A concordância variável no sintagma nominal plural do português reestruturado de Almojarife (São Tomé). *Revista PAPIA*, n.18, p.23-43, 2008.
- _____. *A concordância plural variável no sintagma nominal do português reestruturado da comunidade de Almojarife, São Tomé*. Tese (Doutorado) – Universidade de Macau, Macau, 2010.
- _____. Variável extralinguística escolaridade: influência na marcação plural do sintagma nominal do português reestruturado de Almojarife, São Tomé. *Revista PAPIA*, v.22, n.1, p.41-76, 2012.
- _____. Uso variável do artigo definido no português da comunidade de Almojarife, São Tomé. *Journal of Ibero-Romance Creoles*, n.9, p. 358-389, 2019.
- GONÇALVES, Rita; HAGEMEIJER, Tjerk. O português num contexto multilíngue: o caso de São Tomé e Príncipe. *Revista Científica da Universidade Eduardo Mondlane*, Série: Letras e Ciências Sociais, v.1, n.1, p.87-107, 2015. Disponível em: <<http://repositorio.ul.pt/bitstream/10451/31032/1/Goncalves%26Hagemeijer2015.pdf>>. Acesso em 22 de dez. 2019.
- GUY, Gregory; ZILLES, Ana. *Sociolinguística Quantitativa*. Instrumental de análise. São Paulo: Parábola, 2007.

- HAGEMEIJER, Tjerk. As línguas de São Tomé e Príncipe. *Revista de crioulos de base lexical portuguesa e espanhola*, Macau, v.1, p.1-27, 2009.
- HAGEMEIJER, Tjerk; GONÇALVES, Rita; AFONSO, Beatriz. Línguas e políticas linguísticas em São Tomé e Príncipe. In: PINTO, P. F.; MELO-PFEIFER, S (Org.). *Políticas linguísticas em português*. Lisboa: Lidel, p. 54-79, 2018.
- HASPELMATH, Martin. Definite articles. In: Michaelis, Susanne Maria & Maurer, Philippe & Haspelmath, Martin & Huber, Magnus (eds.). *The atlas of pidgin and creole language structures*. Oxford: Oxford University Press. Disponível em: < <https://apics-online.info/parameters/28>>. Acesso em: 05 jan. 2021.
- HOUAISS, A.; VILLAR, M. S. *Dicionário Houaiss de Língua Portuguesa*. Elaborado pelo Instituto Antônio Houaiss de Lexicografia e Banco de Dados da Língua Portuguesa S/C Ltda. Rio de Janeiro: Objetiva, 2009.
- LABOV, William. *Padrões sociolinguísticos*. Tradução Marcos Bagno; Marta Scherre; Caroline Cardoso. São Paulo: Parábola Editorial, 2008 [1972].
- LONGOBARDI, Giuseppe. How comparative is semantics? A unified parametric theory of bare nouns and proper names. In: *Natural Language Semantics* 9. Holanda: Kluwer Academic Publishers, 335-369, 2001.
- _____. Reference to individuals, person, and the variety of mapping parameters. In: MÜLLER, Henrik Høeg; KLINGE, Alex (eds.). *Essays on Nominal Determination: From morphology to discourse management*. Amsterdã/Filadélfia: John Benjamins Publishing Company, p. 189–211, 2008.
- LORENZINO, Gerardo Augusto. *Retention and Attrition of Umbundu in São Tomé and Príncipe*. SAGE Open, v. 5, n. 4, p.1-15, 2015.
- LUCCHESI, Dante; BAXTER, Alan. A transmissão linguística irregular. In: LUCCHESI, D., BAXTER, A., and RIBEIRO, I. (Orgs). *O português afro-brasileiro*. Salvador: EDUFBA, p. 101-124, 2009.
- LYONS, Christopher. *Definiteness*. Cambridge: Cambridge University Press, 1999.
- MÜLLER, Ana. Nomes nus e o parâmetro nominal no português brasileiro. *Revista Letras*, [S.l.], v. 58, dez. 2002. ISSN 2236-0999. Disponível em: <<https://revistas.ufpr.br/letras/article/view/18363>>. Acesso em: 08 jan. 2020.

- NEVES, Paloma Moore. *SDs nus na posição de objeto no português afrobrasileiro de Helvécia*. Dissertação de mestrado – Universidade Federal da Bahia, Salvador, 2014.
- _____. *Os Sintagmas Determinantes Nus Objetos [+Específicos] [+Definidos] no Português Afro-Brasileiro de Helvécia*. Tese (Doutorado) – Universidade Federal da Bahia, Salvador, 2019.
- SANKOFF, David; TAGLIAMONTE, Sali; SMITH, Eric. *Goldvarb X: a variable rule application for Macintosh and Windows*. Department of Linguistics, University of Toronto, 2005.
- SANTANA, Natali Gomes de Almeida. *As construções dativas no português de duas comunidades bilíngues de São Tomé (África)*. 110 f. il. 2019. Dissertação (Mestrado) – Universidade Federal da Bahia, Salvador, 2019.
- SCHMITT, Cristina; MUNN, Alan. Against the Nominal Mapping Parameter: Bare nouns in Brazilian Portuguese. In: PIUS TAMANJI, MakoHirotni; HALL, Nancy (Orgs.). *Proceeding of NELS 29*. Amherst, MA: GLSA Publications, p. 339-353, 1999.
- SEIBERT, Gerhard. Le massacre de février 1953 à São Tomé: raison d'être du nationalisme santoméen. *Lusotopie*, v. 4, n. 4, p. 173-191, 1997. Disponível em: <<http://www.lusotopie.sciencespobordeaux.fr/seibert97.pdf>>. Acesso em 20 dez. 2019.
- TAGLIAMONTE, Sali. *Analysing Sociolinguistic Variation: Key Topics in Sociolinguistics*. Cambridge University Press: New York, 2006.
- _____. *Variationist Sociolinguistics: Change, Observation, Interpretation*. Wiley Blackwell: Chichester, 2012.
- _____. *Making Waves: The Story of Sociolinguistic Variationist*. Wiley Blackwell: Chichester, 2016.
- WALL, Albert. The distribution of specific and definite bare nominals in Brazilian Portuguese. In: Kabatek, Johannes; Wall, Albert (orgs.). *Bare Noun Phrases in Romance: theory and (empirical) data*. Amsterdam: John Benjamins, 2013.
- WEINREICH, Uriel; LABOV, Willian; HERZOG, Marvin. *Fundamentos empíricos para uma teoria da mudança linguística*. Tradução de Marcos Bagno. São Paulo: Parábola, 2006.

