



UNIVERSIDADE ESTADUAL DE FEIRA DE SANTANA – UEFS  
DEPARTAMENTO DE LETRAS E ARTES  
PROGRAMA DE PÓS-GRADUAÇÃO EM ESTUDOS  
LINGUÍSTICOS



**PAULA FREITAS DE JESUS TORRES**

**ESTRATÉGIAS DE INDETERMINAÇÃO DO SUJEITO NO  
PORTUGUÊS FALADO EM LUANDA-ANGOLA**

FEIRA DE SANTANA, BA  
2024

**PAULA FREITAS DE JESUS TORRES**

**ESTRATÉGIAS DE INDETERMINAÇÃO DO SUJEITO NO  
PORTUGUÊS FALADO EM LUANDA-ANGOLA**

Tese de Doutorado apresentada ao Programa de Pós-Graduação em Estudos Linguísticos, da Universidade Estadual de Feira de Santana, como requisito parcial para obtenção do Título de Doutor em Estudos Linguísticos.

Orientadora: Profa. Dra. Silvana Silva de Farias Araújo.

FEIRA DE SANTANA, BA  
2024

**Ficha catalográfica - Biblioteca Central Julieta Carteado - UEFS**

Torres, Paula Freitas de Jesus  
T647e Estratégias de indeterminação do sujeito no Português falado em  
Luanda - Angola / Paula Freitas de Jesus Torres.- 2024.  
275f.: il.

Orientadora: Silvana Silva de Freitas Araújo

Tese (doutorado) - Universidade Estadual de Feira de Santana.  
Programa de Pós-Graduação em Estudos Linguísticos, 2024.

1. Indeterminação do sujeito. 2. Língua portuguesa – Variação  
angolana. 3. Sociolinguística. 4. Sócio-história. I. Araújo, Silvana  
Silva de Freitas, orient. II. Universidade Estadual de Feira de Santana.  
Programa de Pós-Graduação em Estudos Linguísticos. III. Título.

CDU: 806.90

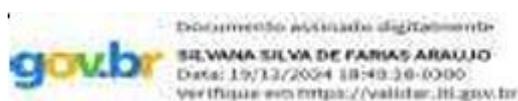
Rejane Maria Rosa Ribeiro – Bibliotecária CRB-5/695

**PAULA FREITAS DE JESUS TORRES**

**ESTRATÉGIAS DE INDETERMINAÇÃO DO SUJEITO NO  
PORTUGUÊS FALADO EM LUANDA-ANGOLA**

Tese apresentada ao Programa de Pós-Graduação em Estudos Linguísticos, da Universidade Estadual de Feira de Santana (UEFS), na linha de pesquisa Variação e Mudança Linguística, como requisito final para a obtenção do título de doutorado.

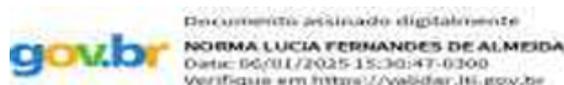
**BANCA EXAMINADORA**



---

Profa. Dra. Silvana Silva de Farias Araújo (Orientadora)

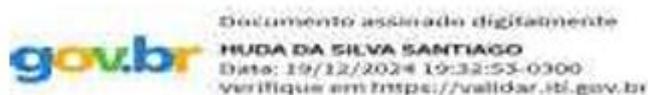
Universidade Estadual de Feira de Santana - UEFS



---

Profa. Dra. Norma Lúcia Fernandes de Almeida (avaliadora interna)

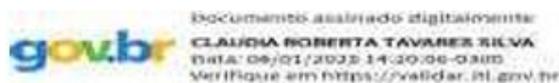
Universidade Estadual de Feira de Santana - UEFS



---

Profa. Dra. Huda da Silva Santiago (avaliadora interna)

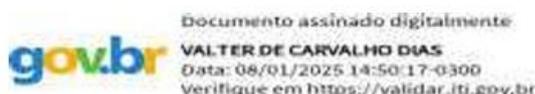
Universidade Estadual de Feira de Santana – UEFS



---

Profa. Dra. Cláudia Roberta Tavares Silva (avaliadora externa)

Universidade Federal de Pernambuco – UFPE



---

Profa. Dr. Valter de Carvalho Dias (avaliador externo)

Instituto Federal de Bahia - IFBA

A todos que constituem minha base de vida, que me dão suporte, que vibram, que riem e que choram comigo (quando necessário).

## AGRADECIMENTOS

A Deus, pela coragem e força para continuar diante das dificuldades encontradas;

A minha família, em especial, meus pais e minha irmã, pelo incentivo, pelo carinho de sempre e, sobretudo, pela compreensão da minha ausência em muitos momentos.

A minha orientadora, Profa. Dra. Silvana Araújo, a quem agradeço enormemente pelos preciosos olhares, correções, pelas orientações não só com o desenvolvimento da tese, mas também pela divisão de experiências durante o meu estágio docência, sugerindo, apoiando as minhas decisões. Sou grata pela força que me deu no momento que mais precisei, pessoalmente e academicamente, incentivando e me apoiando.

Aos meus alunos de Estágio Docência, que me ajudaram a buscar cada vez mais conhecimento, pelas trocas de conhecimento, pelo carinho mesmo em ambiente virtual.

Aos professores do Programa de Pós-graduação em Estudos Linguísticos, pelas contribuições acadêmicas, pelas discussões profícuas e colaborando para o meu crescimento enquanto estudiosa e pesquisadora do Curso do Doutorado em Estudos Linguísticos;

Aos professores da banca de qualificação e defesa, Dra. Norma Lúcia Fernandes de Almeida, Dra. Huda da Silva Santiago, Cláudia Roberta Tavares Silva, Dr. Valter de Carvalho Dias, que com olhar atento e sugestões preciosas, contribuíram grandemente para que o texto fosse aprimorado, cada vez mais.

Aos meus colegas do curso de Doutorado, pelo convívio, pelo compartilhamento das conquistas, pelas risadas ofertadas e aflições durante o decorrer do curso, em especial, Aline e Manoel.

A Jan, um amigo sempre cuidadoso, atencioso e prestativo, desde o início da minha jornada. Foi um anjo que apareceu em meu caminho. Sem palavras para agradecer por toda orientação e carinho dedicados a mim.

A Erick Torres, pelo grande companheirismo, pela compreensão, pela paciência, pela enorme ajuda, principalmente, com a formatação do texto, que, em meio a situações difíceis, sempre esteve ao meu lado e, sobretudo, pelo amor que sempre dedicou a mim.

À Anna Karyna Torres e Moacir Cortês, que sem medir esforços, foram seres de luz que não deixaram as preocupações, incertezas se apossarem de mim. Estenderam, prontamente, a mão, me guiaram e muitas vezes “pausaram” as suas vidas para estarem ao meu lado, me ajudando a seguir em frente.

Ao NELP/UEFS, em nome da coordenadora Dra. Mariana Lacerda, que abriga o *corpus* do projeto de pesquisa “Em busca das raízes do português brasileiro”.

À coordenadora do projeto de pesquisa “Em busca das raízes do português brasileiro”, a professora Silvana Araújo, por sugerir a utilização do material para ampliação dos estudos sobre o português na comunidade de Luanda e aos seus colaboradores, em especial, os alunos de Iniciação Científica, Mateus e Higor, que doaram seu tempo e não mediram esforços para colaborar com esta pesquisa com as transcrições de algumas entrevistas.

Às “As convidadas”, sexteto fantástico, que acompanhou todas as etapas percorridas nos meus estudos de doutoramento, incentivando-me e comemorando comigo todas as conquistas. Não tenho como descrever o acalento nos momentos mais difíceis ao longo do trilhar acadêmico.

Aos meus colegas de profissão, pela força durante o período de estudo;

A todos aqueles que direta ou indiretamente acreditaram em meu potencial e estiveram torcendo pelo meu sucesso acadêmico.

## ADEUS À HORA DA LARGADA

Minha Mãe  
(todas as mães negras  
cujos filhos partiram)  
tu me ensinaste a esperar  
como esperaste nas horas difíceis

Mas a vida  
matou em mim essa mística esperança

Eu já não espero  
sou aquele por quem se espera

Sou eu minha Mãe  
a esperança somos nós  
os teus filhos  
partidos para uma fé que alimenta a vida

Hoje  
somos as crianças nuas das sanzalas do mato  
os garotos sem escola a jogar a bola de trapos  
nos areais ao meio-dia  
somos nós mesmos  
os contratados a queimar vidas nos cafezais  
os homens negros ignorantes  
que devem respeitar o homem branco  
e temer o rico  
somos os teus filhos  
dos bairros de pretos  
além aonde não chega a luz elétrica  
os homens bêbados a cair  
abandonados ao ritmo dum batuque de morte  
teus filhos  
com fome  
com sede com vergonha de te chamarmos Mãe  
com medo de atravessar as ruas  
com medo dos homens  
nós mesmos

Amanhã entoaremos hinos à liberdade  
quando comemormos  
a data da abolição desta escravatura

Nós vamos em busca de luz  
os teus filhos Mãe  
(todas as Mães negras cujos filhos partiram)  
Vão em busca de vida.

(Agostinho Neto)

## RESUMO

Este trabalho investigou as estratégias de indeterminação do sujeito na fala de Luanda (Angola). O trabalho está vinculado ao projeto de pesquisa “Em busca das raízes do Português Brasileiro: estudos morfossintáticos”, em sua fase III, do Núcleo de Estudos em Língua Portuguesa (NELP) da Universidade Estadual de Feira de Santana (UEFS). Esta pesquisa teve também como objetivo contribuir com a agenda de pesquisa do Núcleo supracitado sobre o fenômeno da indeterminação do sujeito na variedade angolana do Português e para isso foram utilizadas 32 entrevistas com falantes de Português como L1, de três faixas etárias em três níveis de escolaridade. Para a realização desse estudo, foram elencadas nove variáveis dependentes. A pesquisa compreendeu variáveis com sujeito lexical preenchido (nós, a gente, você, ele, eu); sem sujeito lexical preenchido, com algumas delas consideradas estratégias canônicas pelas gramáticas tradicionais (o verbo na terceira pessoa do singular acompanhado da partícula “se” / (Ø+V+SE), o verbo na terceira pessoa do plural / (Ø+V3PP), o verbo no infinitivo (Ø+VINF), o verbo na terceira pessoa do singular (Ø+V3PS); as presenças das formas nominais (FN), a exemplo de o homem, o pessoal, entre outros. Foram também levadas em conta, as variáveis linguísticas (grau de indeterminação, função da indeterminação, tipo de frase, tipo de oração, flexão do verbo (SE), tempo e modo verbal, transitividade verbal, tipo de verbo, estrutura do núcleo do predicado, ausência e presença de preposição, concordância com o argumento interno do verbo (SE), posição do argumento interno do verbo, preenchimento do sujeito, mudança/manutenção do referente, forma antecedente/paralelismo, inclusão do falante e as variáveis extralinguísticas (sexo, faixa etária, escolaridade, língua de intercomunicação. Foi usado o método quantitativo com o suporte do programa estatístico *Goldvarb X* e as análises foram feitas à luz da Teoria da Variação e Mudança (LABOV, 1972) bem como foram considerados estudos sobre o fenômeno e os aspectos sócio-históricos marcados por intensos e sucessivos contatos na formação das Variedades Africanas do Português. Os resultados mostraram que houve uma grande variedade de recursos utilizados pelos falantes para indeterminar o sujeito e que eles não se limitam as formas descritas nas GTs, tendo as variáveis internas como as mais favorecedoras. Obteve-se uma menor frequência do *se* como indeterminador nos três níveis de escolaridade, enquanto a forma não preenchida da primeira pessoa do plural (Ø+V1PP) teve um grande destaque, juntamente com a forma pronominalizada (nós). O fenômeno se mostrou passar por um processo de variação estável, mas já apresenta alguns indícios de que uma mudança em progresso poderá ocorrer na fala dos luandenses no futuro. Este trabalho possibilitou uma caracterização do fenômeno variável na comunidade de Luanda com vistas à ampliação de debates com a comunidade científica sobre as realidades sociolinguísticas angolanas.

Palavras-chave: Indeterminação do sujeito. Variedade angolana. Sociolinguística. Sócio-história.

## ABSTRACT

This work investigated the strategies of subject indetermination in the speech of Luanda (Angola). The work is linked to the research project “In search of the roots of Brazilian Portuguese: morphosyntactic studies”, in its phase III, of the Portuguese Language Studies Center (NELP) of the State University of Feira de Santana (UEFS). This research also aimed to contribute to the research agenda of the aforementioned Center on the phenomenon of subject indeterminacy in the Angolan variety of Portuguese. 32 interviews were conducted with speakers of Portuguese as L1, from three age groups and three levels of education. Nine dependent variables were chosen for this study. The research included variables with a completed lexical subject (nós, a gente, você, ele, eu); without a lexical group filled in with some of them considered canonical strategies by traditional grammars (the verb in the third person singular accompanied by the particle “se” / ( $\emptyset$ +V+SE), the verb in the third person plural / ( $\emptyset$ +V3PP), the verb in the infinitive ( $\emptyset$ +VINFIN), the verb in the third person singular ( $\emptyset$ +V3PS); the presence of nominal forms (FN), such as o homem, o pessoal, among others. Linguistic variables were also considered (degree of indeterminacy, function of indeterminacy, type of sentence, type of clause, verb inflection (SE), tense and mood, verbal transitivity, type of verb, structure of the nucleus of the predicate, absence and presence of the verb), argument of the verb (SE), position of the internal argument of the verb, filling in of the subject, change/maintenance of the referent, antecedent form/parallelism, inclusion of the speaker and the socio-historical variables (sex, age group, schooling, language of intercommunication). The quantitative method was used with the support of the Goldvarb X statistical program and the analyses were carried out in the light of LABOV's (1972) Theory of Variation and Change, as well as considering studies on the phenomenon and the socio-historical aspects marked by intense and successive contacts in the formation of African varieties of Portuguese. The results showed a wide variety of resources used by speakers to indeterminate the subject, and that these were not limited to the forms described in the GTs, with internal variables being the most favorable. The use of se as an indeterminate verb was lower across all three levels of education, while the unfilled first-person plural form ( $\emptyset$ +V1PP) was particularly prominent, together with the pronominalized form (we). The phenomenon appeared to be undergoing a stable process of variation, but there are already some indications that a progressive change may occur in the speech of people from Angola in the future. This work has made it possible to characterize the variable phenomenon in the Luanda community with a view to broadening debates with the scientific community on Angolan sociolinguistic realities.

Keywords: Indeterminacy of the subject. Angolan variety. Sociolinguistics. Socio-history.

## RESUMEN

Este estudio investigó las estrategias de indeterminación del sujeto en Luanda (Angola). El trabajo está vinculado al proyecto de investigación «En busca de las raíces del portugués brasileño: estudios morfosintácticos», en su fase III, del Centro de Estudios de la Lengua Portuguesa (NELP) de la Universidad Estatal de Feira de Santana (UEFS). Esta investigación también pretendía contribuir a la agenda de investigación del citado Centro sobre el fenómeno de la indeterminación del sujeto en la variedad angolense del portugués. Se realizaron 32 entrevistas a hablantes de portugués como L1, de tres grupos de edad y tres niveles de educación. Para llevar a cabo este estudio, se eligieron nueve variables dependientes. La investigación incluyó variables con sujeto léxico completado (nós, a gente, você, ele, eu); sin grupo léxico completado con algunas de ellas consideradas estrategias canónicas por las gramáticas tradicionales (el verbo en tercera persona del singular acompañado de la partícula «se» / ( $\emptyset$ +V+SE), el verbo en tercera persona del plural / ( $\emptyset$ +V3PP), el verbo en infinitivo ( $\emptyset$ +VINFIN), el verbo en tercera persona del singular ( $\emptyset$ +V3PS); la presencia de formas nominales (FN), como el hombre, el personal, entre otras. También se tuvieron en cuenta variables lingüísticas (grado de indeterminación, función de la indeterminación, tipo de oración, tipo de cláusula, inflexión verbal (SE), tiempo y modo, transitividad verbal, tipo de verbo, estructura del núcleo del predicado, ausencia y presencia de preposición, concordancia con el argumento interno del verbo (FI), posición del argumento interno del verbo, terminación del sujeto, cambio/mantenimiento del referente, forma del antecedente/paralelismo, inclusión del hablante y variables extralingüísticas (sexo/género, grupo de edad, escolaridad, lengua de intercomunicación). Se utilizó el método cuantitativo con el apoyo del programa estadístico Goldvarb X y los análisis se realizaron a la luz de la Teoría de la Variación y del Cambio de LABOV (1972), además de tener en cuenta los estudios sobre el fenómeno y los aspectos sociohistóricos marcados por intensos y sucesivos contactos en la formación de las variedades africanas del portugués. Los resultados mostraron una amplia variedad de recursos utilizados por los hablantes para indeterminar el sujeto, y que estos no se limitaban a las formas descritas en las GT, siendo las variables internas las más favorables. El uso de *se* como verbo indeterminado fue menor en los tres niveles educativos, mientras que la forma incompleta de la primera persona del plural ( $\emptyset$ +V1PP) fue particularmente prominente, junto con la forma pronominalizada (nós). El fenómeno parecía estar experimentando un proceso estable de variación, pero ya existen indicios de que podría producirse un cambio progresivo en el habla de los luandeses en el futuro. Este trabajo ha permitido caracterizar el fenómeno variable en la comunidad de Luanda con vistas a ampliar los debates con la comunidad científica sobre las realidades sociolingüísticas angolenses.

Palabras clave: Indeterminación del sujeto. Variedad angolense. Sociolingüística. Sociohistoria.

## LISTA DE ABREVIATURAS E SÍMBOLOS

- ALiB – Atlas Linguístico do Brasil
- BA – Bahia
- CPLP - Comunidade dos Países da Língua Portuguesa
- CVC – consoante vogal consoante
- DID – Diálogo entre informante e documentador
- D2 - Diálogo entre dois informantes
- Doc – Documentador
- /ɛ/ - Vogal anterior semiaberta
- EF- Elocução Formal
- FN - Forma nominal
- GNs – Gramáticas Normativas
- GTs - Gramáticas Tradicionais
- IC – Iniciação Científica
- Inf - Informante
- INE – Instituto Nacional de Estatística
- NC – Nasal, Consoante
- NELP – Núcleo de Estudos de Língua Portuguesa
- NUER – Núcleo de Estudos de Identidades e Relações Interétnicas
- NURC – Projeto de Estudo da Norma Urbana Culta no Brasil
- /ɔ/ - Vogal posterior semiaberta
- Oc. - Ocorrências
- Ø + V3SE Verbo na terceira pessoa mais a partícula ‘se’ sem sujeito lexical
- Ø + V3PP - Verbo na terceira pessoa do plural sem sujeito lexical
- Ø + V3PS - Verbo na terceira pessoa do singular sem sujeito lexical
- Ø + VINF - Verbo no infinitivo sem sujeito lexical
- PA - Português Angolano
- PALOPs - Países Africanos de Língua Oficial Portuguesa
- PB – Português Brasileiro
- PVB- Português Vernacular Brasileiro
- PE – Português Europeu
- PHPB – Para a História do Português Brasileiro
- PIDE - Polícia Internacional e de Defesa do Estado
- PL- Português Luandense

PM – Português Moçambicano

P.R. – Peso Relativo

RGPH – Recenseamento Geral da População e Habitação

SN- Sintagma Nominal

SOV - Sujeito Verbo Objeto

STP – São Tomé e Príncipe

SVOX – Sujeito Verbo Objeto e Oblíquos

TLI – Transmissão Linguística Irregular

UEFS – Universidade Estadual de Feira de Santana

VARBRUL – Variable Rule

VARISUL - Variação Linguística na Região Sul do Brasil

VPASSINT - Voz passiva sintética

VPSA - Voz passiva sem agente

## LISTA DE FIGURAS

Figura 1: Principais rotas do tráfico transatlântico .....	41
Figura 2: Esboço de mapa etnológico africano no Brasil .....	46
Figura 3: Modelo de aquisição para situação de contato (TLI) .....	58
Figura 4: Mapa dos países que integram os PALOP .....	63
Figura 5: Línguas da África: troncos linguísticos, segundo Petter .....	81
Figura 6: Mapa nigero-congolês (principais línguas) .....	82
Figura 7: Classificação das zonas geolinguísticas de Guthrie .....	83
Figura 8: Grupos linguísticos do tronco afro-asiático .....	84
Figura 9: Classificação do tronco Nilo-saariano .....	86
Figura 10: A Origem e as direções migratórias dos povos bantu .....	88
Figura 11: Mapa etnolinguístico de Angola .....	92
Figura 12: Localização de Angola e países fronteiriços .....	94
Figura 13: Municípios, comunas e províncias angolanas (2014) .....	95
Figura 14: Línguas faladas pela população com dois ou mais anos pela área de residência (Zona rural, zona urbana e Angola) .....	97
Figura 15: Localização cidade de Luanda (Angola) no território africano .....	158
Figura 16: As 18 províncias de Angola, 2016 .....	160
Figura 17: <i>Continuum</i> (graus de indeterminação) .....	176

## LISTA DE GRÁFICOS

Gráfico 1 - Distribuição geral das 15 variantes .....	213
Gráfico 2 - Distribuição geral das 06 estratégias descartadas .....	215
Gráfico 3 - Distribuição geral das 09 estratégias indeterminadoras .....	219
Gráfico 4 - Distribuição geral das 07 estratégias indeterminadoras .....	225

## LISTA DE QUADROS

Quadro 1: Áreas de procedências dos cativos .....	39
Quadro 2: Línguas africanas trazidas pelo tráfico ao Brasil (Século XVI ao XIX) .....	43
Quadro 3: Distribuição demográfica no Brasil do século XVI-XIX .....	44
Quadro 4: População escravizada .....	45
Quadro 5: Traços linguísticos das línguas bantas.....	89
Quadro 6: Principais grupos de línguas, povos e respectivas áreas dos povos angolanos.....	90
Quadro 7: Províncias de Angola em relação à população .....	95
Quadro 8: Resumo do resultado de pesquisas desenvolvidas no Brasil e em Angola sobre o fenômeno da indeterminação do sujeito. ....	137
Quadro 9: Blocos de Interesse (entrevistas) .....	155
Quadro 10: Estratificação da Amostra das entrevistas do Projeto Em busca das raízes do Português Brasileiro .....	156
Quadro 11: Participantes Corpus Luanda (Português L1) .....	161
Quadro 12: Visão resumida dos Informantes (Português L1) .....	162
Quadro 13: Variável dependente e seus respectivos fatores .....	163
Quadro 14: Variável dependente – formas nominais.....	164
Quadro 15: Variáveis controladas.....	165
Quadro 16 - Pronominais X não-pronominais com dados apenas de “nós” .....	249
Quadro 17 - Pronominais X não-pronominais com dados de “nós”+ “Ø+V1PP” .....	250

## LISTA DE TABELAS

Tabela 1 - Distribuição geral das estratégias de indeterminação do sujeito.....	211
Tabela 2 - Distribuição geral das 06 variantes descartadas de indeterminação do sujeito encontradas no <i>corpus</i> .....	214
Tabela 3 - Distribuição geral das 09 variantes de indeterminação do sujeito.....	217
Tabela 4 – Comportamento da estratégia "ØVIPPP" diante das demais variantes. ....	221
Tabela 5 - Influência do ØVIPPP no uso da estratégia <i>nós</i> indeterminador.....	223
Tabela 6 - Influência do <i>nós</i> no uso da estratégia ØVIPPP indeterminador.....	223
Tabela 7- Frequência das estratégias de indeterminação do sujeito dentre os dados analisados .....	224
Tabela 8 - Influência da escolaridade no uso da estratégia <i>nós</i> indeterminador .....	227
Tabela 9 - Influência da função da indeterminação no uso da estratégia <i>nós</i> indeterminador .....	228
Tabela 10 - Influência do paralelismo no uso da estratégia <i>nós</i> indeterminador .....	229
Tabela 11 - Influência da mudança/manutenção do referente no uso da estratégia <i>nós</i> indeterminador.....	230
Tabela 12 - Influência do paralelismo no uso da estratégia <i>nós</i> indeterminador .....	232
Tabela 13 - Influência do preenchimento do sujeito no uso da estratégia <i>nós</i> indeterminador .....	233
Tabela 14 - Influência do tempo e modo verbal no uso da estratégia <i>nós</i> indeterminador ...	233
Tabela 15 - Influência do tipo de verbo no uso da estratégia <i>nós</i> indeterminador.....	234
Tabela 16 - Influência do grau de indeterminação no uso da estratégia <i>nós</i> indeterminador .....	236
Tabela 17 - Influência do tempo e modo verbal no uso da estratégia <i>nós</i> indeterminador ...	237
Tabela 18 - Influência do tipo de oração no uso da estratégia <i>nós</i> indeterminador.....	237
Tabela 19 - Influência do paralelismo no uso da estratégia <i>nós</i> indeterminador .....	239
Tabela 20 - Influência do paralelismo no uso da estratégia <i>nós</i> indeterminador .....	242
Tabela 21 - Influência da língua de intercomunicação no uso da estratégia <i>nós</i> indeterminador.....	243
Tabela 22 - Influência da escolaridade no uso da estratégia <i>nós</i> indeterminador .....	244
Tabela 23 - Influência do grau de indeterminação no uso da estratégia <i>formas nominais</i> em relação às demais estratégias .....	246
Tabela 24 - Influência da inclusão do falante no uso da estratégia <i>formas nominais</i> em relação às demais estratégias.....	247
Tabela 25 - Influência da mudança/manutenção do referente no uso da estratégia <i>formas nominais</i> em relação às demais estratégias .....	247

Tabela 26 - Influência do tempo e modo verbal no uso da estratégia <i>formas nominais</i> em relação às demais estratégias .....	248
Tabela 27 – Comportamento da estratégia “nós” diante das demais variantes.....	251

## SUMÁRIO

<b>1 INTRODUÇÃO</b> .....	<b>20</b>
<b>2 VARIEDADES NÃO EUROPEIAS DA LÍNGUA PORTUGUESA: BRASILEIRAS E AFRICANAS</b> .....	<b>27</b>
2.1 LÍNGUA PORTUGUESA: FORMAÇÃO E TRANSPLANTAÇÃO .....	27
2.2 VARIEDADE DO PORTUGUÊS NO CPLP/PALOP: RELAÇÕES ÁFRICA-BRASIL	28
2.2.1 A diáspora africana .....	<b>29</b>
2.2.1.1 Dos tumbeiros às <i>plantations</i> .....	29
2.2.1.2 Reflexões sobre a língua portuguesa em contato com os povos autóctones e africanos no Brasil colonial e independente .....	33
<b>2.2.2 As línguas bantu e a emergência de uma variedade brasileira do português.</b>	<b>39</b>
2.2.2.1 Deriva secular das línguas românicas .....	49
2.2.2.2 Contato linguístico .....	53
2.3 VARIEDADES DO PORTUGUÊS EM ÁFRICA: PALOP .....	61
<b>3 VARIEDADE ANGOLANA DA LÍNGUA PORTUGUESA: ASPECTOS SÓCIO-HISTÓRICOS E LINGUÍSTICOS</b> .....	<b>79</b>
3.1 CONhecendo os troncos linguísticos em África: as línguas e os povos.....	80
<b>3.1.1 As línguas do grupo bantu</b> .....	<b>87</b>
3.1.1.1 As línguas bantu e os grupos étnicos em Angola.....	90
3.2 REPÚBLICA DE ANGOLA.....	94
<b>3.2.1 História social de Angola</b> .....	<b>98</b>
<b>3.2.2 Formação do Português Angolano</b> .....	<b>102</b>
<b>3.2.3. Relações oralidade/escrita nas culturas africanas</b> .....	<b>104</b>
<b>3.2.4 Multilinguismo em Angola: línguas locais e língua portuguesa</b> .....	<b>106</b>
<b>3.2.5 Ensino da língua portuguesa hoje em Angola e as variedades endógenas</b> ....	<b>110</b>
<b>4 O FENÔMENO LINGUÍSTICO: INDETERMINAÇÃO DO SUJEITO</b> .....	<b>114</b>
4.1 SUJEITO: ALGUMAS CONCEPÇÕES.....	114
4.2 A INDETERMINAÇÃO DO SUJEITO.....	115
<b>4.2.1 Abordagens tradicionais para a indeterminação do sujeito</b> .....	<b>115</b>

<b>4.2.2 Outras concepções: determinação, indeterminação e indefinição para o estudo do sujeito</b> .....	<b>119</b>
<b>4.2.3 Estudos sobre a indeterminação do sujeito Brasil/Angola</b> .....	<b>121</b>
4.2.3.1 Estudos Linguísticas em amostras de fala no Brasil .....	122
4.2.3.2 Estudos Linguísticas em amostras escritas no Brasil.....	134
4.2.3.3 Estudos Linguísticas em amostras de fala em Angola .....	135
<b>5 PRESSUPOSTOS TEÓRICO-METODOLÓGICOS</b> .....	<b>143</b>
5.1 SOCIOLINGÜÍSTICA LABOVIANA: TEORIA DA VARIAÇÃO E MUDANÇA LINGÜÍSTICA.....	143
5.2 O MÉTODO DA SOCIOLINGÜÍSTICA QUANTITATIVA LABOVIANA .....	148
<b>5.2.1 A ferramenta computacional <i>Goldvarb X</i></b> .....	<b>149</b>
<b>5.2.2 Etapas da pesquisa</b> .....	<b>150</b>
5.3 O <i>CORPUS</i> .....	152
<b>5.3.1 Projeto “Em busca das raízes do Português Brasileiro”</b> .....	<b>152</b>
<b>5.3.2 As entrevistas</b> .....	<b>153</b>
<b>5.3.3 A transcrição</b> .....	<b>157</b>
<b>5.3.4 A comunidade: Luanda</b> .....	<b>157</b>
<b>5.3.5 Os participantes da pesquisa</b> .....	<b>160</b>
<b>5.3.6 Seleção das variáveis</b> .....	<b>163</b>
5.4 VARIÁVEIS CONTROLADAS .....	169
<b>5.4.1 Variável dependente</b> .....	<b>171</b>
5.4.1.1 Verbo na terceira pessoa do singular acompanhado da partícula “se” / Ø + V+SE.....	172
5.4.1.2 Verbo na terceira pessoa do plural / Ø + V3PP .....	173
5.4.1.3 Verbo na terceira pessoa do singular / Ø+V3PS.....	174
5.4.1.4 Verbo no infinitivo impessoal / Ø + VINF .....	175
5.4.1.5 Verbo na voz passiva sem agente (VPSA).....	176
5.4.1.6 Formas com o sujeito lexical preenchido.....	177
5.4.1.6.1 Nós .....	178

5.4.1.6.2 A gente.....	179
5.4.1.6.3 Você.....	179
5.4.1.6.4 Eles .....	180
5.4.1.6.5 Eu.....	181
5.4.1.6.6 Formas nominais.....	181
<b>5.4.2 Variáveis Independentes.....</b>	<b>182</b>
5.4.2.1 Variáveis Linguísticas.....	183
5.4.2.1.1 Grau de Indeterminação do contexto.....	183
5.4.2.1.2 Função da Indeterminação.....	185
5.4.2.1.3 Tipo de frase.....	186
5.4.2.1.4 Tipo de oração.....	187
5.4.2.1.5 Flexão verbal (SE).....	187
5.4.2.1.6 Tempo e modo verbal.....	188
5.4.2.1.7 Transitividade verbal.....	190
5.4.2.1.8 Tipo de verbo.....	190
5.4.2.1.9 Estrutura do núcleo do predicado.....	191
5.4.2.1.10 Ausência ou presença de preposição.....	192
5.4.2.1.11 Concordância com o argumento interno (partícula <i>SE</i> ).....	193
5.4.2.1.12 Posição do argumento interno do verbo.....	194
5.4.2.1.13 Preenchimento do sujeito.....	195
5.4.2.1.14 Mudança e Manutenção do referente (Paralelismo).....	195
5.4.2.1.15 Forma Antecedente (Paralelismo).....	197
5.4.2.1.16 Inclusão do Falante.....	198
5.4.2.2 Variáveis Extralinguísticas.....	198
5.4.2.2.1 Sexo.....	199
5.4.2.2.2 Faixa Etária.....	201
5.4.2.2.3 Escolaridade.....	202
5.4.2.2.4 Língua de Intercomunicação.....	204

<b>6 AS ANÁLISES E INTERPRETAÇÕES SOCIAIS E LINGUÍSTICAS DOS DADOS DE FALA DO POVO DE LUANDA .....</b>	<b>206</b>
6.1 VOZES DE LUANDA EM ANÁLISE: O SOCIAL E LINGUÍSTICO SE ENCONTRAM .....	206
6.2 VOZES DE LUANDA: OS DADOS LINGUÍSTICOS EM FOCO .....	211
<b>6.2.1 Rodada da estratégia de Indeterminação Nós versus demais variantes.....</b>	<b>223</b>
<b>6.2.2 Nós versus as demais .....</b>	<b>225</b>
<b>6.2.3 Nós versus A gente.....</b>	<b>226</b>
<b>6.2.4 Nós versus Você .....</b>	<b>231</b>
<b>6.2.5 Nós versus Formas nominais .....</b>	<b>235</b>
<b>6.2.6 Nós versus Ø+V3PS.....</b>	<b>238</b>
<b>6.2.7 Nós versus Ø+V3PP.....</b>	<b>240</b>
<b>6.2.8 Nós versus Ø+V+Se .....</b>	<b>241</b>
<b>6.2.9 Nós versus Ø+VINF.....</b>	<b>242</b>
6.3 FORMAS NOMINAIS.....	244
6.4 FORMAS PRONOMINAIS X NÃO PRONOMINAIS.....	249
<b>7 CONSIDERAÇÕES FINAIS.....</b>	<b>254</b>
<b>REFERÊNCIAS.....</b>	<b>260</b>
<b>APÊNDICE A – TABELA DE CODIFICAÇÃO .....</b>	<b>273</b>

## 1 INTRODUÇÃO

O surgimento da Sociolinguística, na década de 1960, propiciou uma mudança no estudo da língua, principalmente pelo fato de dar maior visibilidade à variedade falada: constatou-se que a língua varia e as diferentes formas empregadas pelos falantes devem ser também consideradas legítimas e entendidas como um produto sociocultural, interativo e coletivo em uma dada comunidade. Assim, é de grande importância considerar o papel da sociedade nos processos de variação e mudança pelos quais a língua passa, seja ela oral ou escrita, correlacionando aspectos linguísticos e sociais no trato da heterogeneidade, que é sistemática e ordenada. É necessário observar os contextos semântico-pragmático e sócio-histórico, nos quais são produzidos os usos linguísticos, tornando, assim, possível uma análise mais coerente da língua, sem reduzir a análise linguística apenas a seu nível sintático.

As análises sobre a indeterminação do sujeito, fenômeno investigado nesta tese, apresentadas pelas gramáticas normativas, não dão conta de explicar, de forma satisfatória, várias questões. Para Milanez (1982), algumas delas são o conceito de indeterminação, impessoalização, indefinição e, de igual forma, a ausência de análises concomitantes nos níveis sintáticos, semânticos e pragmáticos. Outro ponto importante é a não consideração de recursos utilizados na modalidade oral (as variantes). Para essa autora, a indeterminação do sujeito, assim como qualquer fenômeno linguístico, deve ser vista como um fenômeno dependente de contexto para que seja possível uma interpretação mais coerente. As questões apresentadas reforçam a limitação das gramáticas e dos manuais quanto suas definições e seus registros, limitação essa que a faz se distanciar do desempenho linguístico real dos falantes.

Há de considerar que ainda não existe consenso para definir de forma clara e satisfatória a indeterminação do sujeito no português brasileiro (doravante PB), pelas gramáticas tradicionais. Enquanto a maioria dos manuais apontam basicamente duas formas para o sujeito indeterminado, *verbo na terceira pessoa do plural* e o pronome *se ligado a verbos na terceira pessoa do singular*, os resultados de pesquisas já realizadas mostraram que os falantes, além das formas já legitimadas,  $\emptyset+3PP$  e  $\emptyset+V+SE$ , utilizam também  $\emptyset+3PS$ , o  $\emptyset+VINF$  e outras formas pronominais como *eu, tu, você, a gente, nós, eles, formas nominais (FN)* para indeterminar o sujeito, além de várias outras. As construções<sup>1</sup> (1) e (2), hipotéticas,

---

<sup>1</sup> O papel do contexto é muito importante para ajudar na compreensão dos exemplos e no entendimento para definir se o sujeito é indeterminado ou não. Embora haja a necessidade de um contexto maior, um contexto interno nos exemplos hipotéticos, optou-se por mostrar as estratégias encontradas, como exemplificação de possíveis

exemplificam esses usos consagrados pelos compêndios gramaticais. Já as construções (3) a (9), também hipotéticas, exemplificam algumas outras formas para indeterminar o sujeito das orações, com estratégias já documentadas na literatura.

- (1) **Contaram-me** um segredo.
- (2) **Descobriu-se** o segredo.
- (3) **A gente** acredita em tudo.
- (4) **Eles** não acreditam em nada disso.
- (5) **Nós** sofremos com as consequências.
- (6) É triste quando **o indivíduo** não segue as regras.
- (7) **A pessoa** chama os outros e veja no que dá.
- (8) **Ø** corre, **Ø** grita, **Ø** chora demais.
- (9) **Diz** que o livro é muito importante.

É importante destacar que há uma grande variedade de estudos no PB, como os trabalhos de Duarte (1995, 2010) que apontam formas distintas para o uso desse fenômeno, as quais ainda não foram legitimadas pelas gramáticas tradicionais, como pode ser verificado nos trabalhos de Milanez (1982) com dados do projeto NURC/SP; Rollemberg (1991) e Carvalho (2010), ambos com dados coletados na cidade de Salvador; Santana (2006) ao estudar o fenômeno no português rural do semiárido baiano; Setti (1997), com as capitais do sul do Brasil; Godoy (1999); Assunção (2012); Assunção e Almeida (2016), Souza (2014) e Souza e Oliveira (2014), na cidade de Feira de Santana; Cunha (1993), Dias (2017), entre outros. Na variedade angolana, os estudos ainda são muito insipientes, mas pode-se destacar os estudos de Teixeira (2008) sobre o pronome *você* no português da capital Luanda, com dados coletados pela própria pesquisadora; Teixeira e Almeida (2011) que comparam o fenômeno da indeterminação do sujeito no português angolano com o português brasileiro.

Diante das discussões já feitas com dados do PB no tocante às visões incoerentes e indefinidas próprias das gramáticas tradicionais, é necessário continuar investigando e analisando o uso real que os falantes fazem da sua língua não só no PB, mas também em outras variedades do português, a exemplo do português falado em Luanda (variedade focalizada nesta tese) quanto à indeterminação do sujeito, que também podem ir de encontro às formas

---

realizações para indeterminar o sujeito. Vale ainda ressaltar que, em alguns casos, no decorrer das análises encontradas no *corpus*, alguns contextos internos não vão aparecer visto que a cadeia discursiva nem sempre é curta.

extremamente limitadas pelas GTs. Desse modo, é imprescindível responder a algumas questões nesta pesquisa:

- a) quais são os recursos de indeterminação do sujeito mais utilizados pelos falantes luandenses, já que esse fenômeno está sendo largamente verificado no PB em modalidades oral e escrita em seus usos vernacular e culto?
- b) os possíveis resultados verificados em Luanda apresentarão similaridades com os resultados já encontrados na variedade popular oral do Português Brasileiro?
- c) quais contextos favorecem mais o uso das formas não legitimadas pelas gramáticas tradicionais, os linguísticos ou os extralinguísticos?
- d) a indeterminação do sujeito se caracteriza por um processo de variação estável ou de mudança em progresso na fala da capital de Angola?

Essas questões são postas partindo do pressuposto de que o falar de Luanda atual, assim como já se percebe no Português Brasileiro, pode estar passando por um processo de variação/mudança em relação ao Português Europeu. Portanto, as hipóteses a serem testadas nesta pesquisa sobre a indeterminação do sujeito são:

- a) Sabendo que o contato entre línguas e povos distintos favorece o aparecimento de novas variedades linguísticas, é possível que o PA esteja passando pela mesma variabilidade/mudança no sistema, no que concerne a indeterminação do sujeito, apresentando semelhanças com o que já se verifica no PB, em especial por falantes da norma popular;
- b) As estratégias utilizadas pelos sujeitos luandenses não se limitam às formas de indeterminação do sujeito prescritas pelas gramáticas normativas, gramáticas tradicionais;
- c) O pronome *se* como indeterminador do sujeito deve estar em desuso na fala daqueles com graus de escolaridade mais baixos, visto que essa é uma estratégia aprendida na escola, assim como as variantes mais inovadoras sejam mais utilizadas por falantes menos escolarizados;
- d) O princípio do paralelismo sintático-discursivo pode ser também uma das variáveis linguísticas mais favorecedoras no uso de diferentes estratégias de indeterminação do sujeito no PA assim como acontece com o PB.

O objetivo deste trabalho foi verificar qual é a realidade linguística do português falado em Luanda, especificamente no uso do fenômeno estudado (indeterminação do sujeito), a fim de fornecer dados que possibilitem compreender o entendimento dos contatos linguísticos e dos processos de mudanças ocorridas nas áreas em que a língua portuguesa se expandiu. Como objetivos específicos, se propõe:

- Identificar os contextos linguísticos (tempo e modo verbal, tipo de oração, tipo de verbo e preenchimento ou não do sujeito, grau de indeterminação entre outros) e extralinguísticos (sexo, faixa etária, escolaridade, língua de intercomunicação) que condicionam o uso das variantes apresentadas pela gramática tradicional e outras formas que podem ser utilizadas como estratégias pelos falantes da indeterminação do sujeito;
- Discutir sobre os estudos linguísticos em relação ao fenômeno da indeterminação do sujeito no Português Angolano, examinando se os fatores condicionantes (linguísticos e sócio-históricos) atuam de maneira semelhante em relação ao mesmo fenômeno linguístico na variedade brasileira;
- Comparar os resultados obtidos neste estudo com os alcançados por outros autores que consideraram dados do PB, ampliando o debate sobre a realidade sociolinguística brasileira e angolana em seu aspecto formativo.

Vários trabalhos de cunho sociolinguístico já foram realizados e permitem traçar um perfil dos diferentes falares em diversas regiões. Dentre esses estudos já realizados, em especial sobre a formação do Português Brasileiro, as análises comparativas concentram-se mais nas variedades brasileira e europeia e isso limita a ampliação de discussões sobre as influências de diferentes contatos linguísticos na formação de outras variedades do português ou de línguas crioulas. Em razão disso, faz-se necessário que os estudos também se voltem para outras variedades do Português, especificamente para variedades faladas em outras ex-colônias portuguesas, a fim de compreender melhor se essas variedades apresentam similaridades com o Português Brasileiro devido às causas exógenas, como é o caso do intenso contato linguístico do povo português na África e no Brasil, e situações de transmissão linguística irregular, embora cada uma delas com suas peculiaridades.

A necessidade de ampliação dos estudos variacionistas com dados coletados em diferentes países onde se fala o Português é defendida por Petter (2007), ao afirmar que estudar variedades africanas do português ajuda a fornecer dados que possibilitam o entendimento

acerca dos contatos linguísticos e dos processos de mudanças ocorridos nas áreas em que a língua portuguesa se expandiu, investigando o peso do contato entre línguas diversas e ininteligíveis. A autora aponta que, cada vez mais, se tem um distanciamento da norma europeia, pois já há indícios de semelhanças entre as variedades não europeias, como as angolanas, as moçambicanas e as brasileiras, aventando, assim, a hipótese de haver um *continuum* afro-brasileiro do Português, haja vista que os contatos entre as línguas africanas e o português produziram consequências linguísticas semelhantes por todas compartilharem línguas do tronco bantu, além de resultarem fortemente de um processo de expansão da Língua Portuguesa, em decorrência das expansões ultramarinas, em um contexto de colonização e provenientes de contatos entre línguas africanas e do português.

Nessa perspectiva, o estudo sobre o fenômeno da indeterminação do sujeito em variedades do português ainda não tão conhecida, como é o caso do Português Angolano, poderá propiciar o conhecimento não só em termos linguísticos dessa outra variedade, mas também dos seus aspectos sócio-históricos que permitam a ampliação de investigações das próprias origens do Português Brasileiro.

Ademais, acredita-se que o desenvolvimento de uma pesquisa como esta tenha grande importância para outros estudos sobre as variedades do português, sejam elas brasileira, europeia, sejam africanas, ao fornecer contribuições significativas no trato do fenômeno morfossintático da indeterminação do sujeito, estudo este que se mostra incipiente na capital angolana. Conhecer as formas possíveis das variantes utilizadas pelos angolanos em seu uso real de produção favorece o conhecimento dos aspectos divergentes e, principalmente, dos convergentes entre essas variedades, baseando-se nas situações de contatos maciços entre seus povos que vieram a contribuir para a formação da língua portuguesa, respeitando as suas peculiaridades, suas histórias, suas diversidades.

Diante do que já foi exposto, o presente trabalho está organizado em sete seções. Na Seção 2, foca-se em discussões sobre questões relacionadas às situações de contato, à variedade não europeia de língua portuguesa, em especial, a brasileira, relacionando a ligação com os povos africanos mesmo antes de aportarem em solo americano. Inicialmente, são apresentadas essas relações entre Brasil – África, focando as influências das línguas africanas no português que emergia. São também discutidas propostas teóricas para o contato linguístico, como a deriva românica, à transmissão linguística irregular e à pidginização/crioulização. Nessa seção, são ainda descritas, outras variedades do Português em África fazendo uma correlação com aspectos que foram discutidos sobre a variedade brasileira, muitas das suas similaridades e

aspectos distintivos, em consequência dos intensos contatos existentes entre os portugueses e os povos africanos.

Na Seção 3, apresenta-se a variedade africana da língua portuguesa, com foco na variedade angolana. Discutem-se os aspectos sócio-históricos e linguísticos de Angola, a formação do português e a grande diversidade de línguas em seu território, oriundo de diferentes grupos linguísticos, em especial os do grupo bantu. São discutidas, também, as políticas para a implantação da língua portuguesa em Angola, as línguas locais e a situação de multilinguismo intenso em que o país se encontra.

Na seção 4, discorre-se sobre as definições dadas pelas gramáticas tradicionais brasileiras/portuguesas sobre os distintos pontos de vista para definições do sujeito. O mesmo é feito sobre a indeterminação do sujeito, sem perder de vista as concepções não tradicionais, as concepções que trabalham com a temática, mas apontam outros caminhos, outras estratégias de indeterminação do sujeito que podem ser consideradas legítimas, que são produzidas pelos falantes, uma vez que se caracteriza como um fenômeno variável da língua. Desse modo, também são discutidos alguns conceitos do que pode ser considerado determinado ou não e outras definições que se julgarem necessárias. Nessa seção, ainda são apresentados trabalhos já realizados sobre o fenômeno, especialmente no território brasileiro, e seus resultados serão discutidos.

Na seção 5, são descritos os pressupostos teórico-metodológicos da Teoria da Variação e Mudança. Nessa seção, é descrito o *corpus* que foi utilizado do Projeto “Em busca das raízes do Português Brasileiro”, sediado no NELP/UEFS (Teixeira; Araujo, 2017). Apresentam-se informações do *corpus* investigado (Luanda), não somente, seus participantes, como, as entrevistas foram realizadas e transcritas com vistas à codificação dos dados, mas também o método quantitativo de análise e instrumentos utilizados a fim de possibilitar uma análise que será realizada na seção seguinte. Também são apresentadas as variáveis controladas nessa pesquisa, a variável dependente e as independentes (linguísticas e extralinguísticas) com base em estudos prévios sobre o fenômeno.

Na seção 6, apresentam-se os dados com base nas variáveis descritas para a investigação na seção anterior, os resultados das análises e as suas interpretações, tendo como base as variáveis selecionadas e apontadas pelo programa estatístico como mais favorecedoras. São feitas análises para conhecer a real situação linguística da cidade de Luanda em relação ao tema desse estudo. São mostradas as análises binárias. É, também, nessa seção, que são descritas as estratégias de indeterminação do sujeito mais utilizadas na comunidade investigada, assim, valores percentuais e de pesos relativos são tomados como parâmetros para tal análise.

Na seção 7, encontram-se conclusões tomadas após análises dos dados sobre o fenômeno, na comunidade de Luanda, bem como futuros passos que podem ser realizados para que a situação sociolinguística da capital de Luanda possa ser mais bem apreendida, ampliando-se para outros estudos de variedades africanas.

Espera-se que este trabalho possa agregar e fortalecer as pesquisas já feitas no Projeto “Em busca das raízes do português brasileiro”, fase 3, do NELP/UEFS do ponto de vista da morfossintaxe com outras agendas de pesquisa da equipe na tentativa de conhecer os motivos pelos quais a variedade angolana vernacular se assemelha ou se distancia em relação à brasileira. Pretende-se, obter uma melhor compreensão do papel das línguas africanas na formação de variedades do português nos âmbitos linguísticos e sócio-históricos, tanto do PB quanto do PA, já que ambas foram frutos de situações intensas de contato devido ao processo de colonização portuguesa escravagista. Busca-se também a caracterização do fenômeno variável na comunidade de Luanda com vistas à ampliação de debates com a comunidade científica sobre as realidades sociolinguísticas angolana e brasileira.

## 2 VARIEDADES NÃO EUROPEIAS DA LÍNGUA PORTUGUESA: BRASILEIRAS E AFRICANAS

*Ó árvore do esquecimento;  
 Símbolo de todo o meu tormento.  
 Por que nasceste à beira mar?  
 Por que nasceste?  
 Cruzei os mares sem te ver, meu Benin;  
 Fui arrancado do teu colo, minha mãe-terra;  
 Quiseram roubar-me as lembranças de ti,  
 Violentaram meu corpo e minha dignidade,  
 Mas minha alma e meu coração estão intactos? deixei-os aí.  
 Só tocaram o que podiam ver;  
 O que não podiam, permaneceu puro.  
 Quantas vezes quis te reencontrar, chão pátrio;  
 Terra dos meus ancestrais, lar dos meus pais,  
 Para aonde retorno todas as noites quando o corpo descansa.  
 (Banzo, Cleber Menezes)*

Nesta seção, discutem-se questões sobre a transplantação da Língua Portuguesa e sobre aspectos relacionados às situações de contato linguístico, a exemplo das propostas teóricas da deriva românica, da transmissão linguística irregular e da pidginização/crioulização. Foca-se em alguns dos principais acontecimentos sociais, históricos e geográficos que ocorreram no processo de dominação de novos territórios pelo povo português nos continentes americano e africano. Desse modo, abordam-se a formação sócio-histórica e a constituição de variedades não europeias do Português, dando ênfase ao significativo papel do elemento negro e bantu, primeiramente, considerando o contexto do Brasil e, depois, o contexto de países africanos.

### 2.1 LÍNGUA PORTUGUESA: FORMAÇÃO E TRANSPLANTAÇÃO

A língua portuguesa hoje é uma das línguas mais faladas do mundo, sendo o Brasil, o país com o maior número de falantes e o único, nas Américas, onde o português é língua oficial, justamente por ter sido o único com colonização portuguesa. Para melhor compreender o PB, é preciso voltar no tempo para a formação do PE. Fazendo algumas considerações sobre a formação da língua portuguesa na Europa, percebe-se que esse processo foi resultado de influências de várias línguas de outros povos da Península Ibérica e antecede muito a invasão do Brasil pelos portugueses em 1500. As invasões germânicas e árabes no território português deixaram marcas na língua, mas o Latim foi, sem dúvida, um dos grandes influenciadores para o nascimento do idioma Português. À medida que Roma conquistava novos povos, obrigava o uso do latim como língua oficial por toda a extensão do Império Romano. Já, quanto à formação

das novas variedades da Língua Portuguesa, tanto no Brasil como em países africanos, não há como ignorar a participação dos povos autóctones (os originários no Brasil e os africanos em África) e dos traficados, estes, notadamente, os africanos que foram trazidos para o Brasil.

Na origem da Língua Portuguesa na Europa, como os povos conquistados eram muitos e falavam línguas diferenciadas, o latim vulgar, modalidade imposta aos vencidos, foi também sofrendo modificações nas distintas regiões, o que veio a gerar, posteriormente, as diferentes línguas neolatinas, como o catalão, o castelhano e o galego-português. Esse último, dialeto falado na parte ocidental da Península Ibérica, à proporção que foi se adentrando para o sul, se alterou e predominou na região. Com a separação e independência do Reino português, o galego e o português distanciaram-se de modo que o português se tornou a língua, parte da construção e expansão do Império Português. Portanto, é necessário ressaltar que o português foi um desenvolvimento do latim vulgar nos seus variados aspectos como, por exemplo, a fonética, morfologia, léxico e sintaxe.

Como os ventos que moveram os navios nas tenebrosas expansões e invasões, o português se espalhou pelo mundo, variando em suas colônias. Sons diversos deram formas aos mesmos vocábulos, uma língua imposta começa a exprimir as angústias e anseios de povos dominados, separados geograficamente, mas unidos por dores análogas. Seja nos países da África ou em solo brasileiro, a língua dos colonizadores portugueses, da mesma forma que representava uma mudança imposta, a dominação, o subjugo dos conquistados, também sofria as mudanças, as interferências e a presença dos subtraídos que, paulatinamente, permearam a língua com suas próprias vozes.

## 2.2 VARIEDADE DO PORTUGUÊS NO CPLP/PALOP: RELAÇÕES ÁFRICA-BRASIL

Antes restrita à área da Península Ibérica, a partir do século XV, a conquista por novos territórios pelos portugueses acabou por impulsionar também a expansão da língua portuguesa para outros continentes, a exemplo do americano e do africano. Na América, o contato da língua Portuguesa com as do grupo banto aconteceu devido a vinda forçada de milhares de africanos para o Brasil. Já, nos países africanos, esse contato se deu com os africanos que lá já estavam. No Brasil, o português foi ganhando identidades indígenas, negras e de imigrantes europeus até ganhar o ritmo inconfundível de hoje.

Na seção seguinte, é discutido um pouco das relações dos povos africanos com o povo português e sua língua bem como a participação no processo de formação da variedade nacional.

## 2.2.1 A diáspora africana

### 2.2.1.1 Dos tumbeiros às *plantations*

A escravização de africanos provocou um deslocamento muito intenso das populações da África e a transposição de elementos da cultura africana, seus estilos de vida, suas práticas religiosas, suas línguas para terras que estavam sendo ocupadas. Quando a população africana (composta de vários grupos étnicos e linguísticos) era colocada nos navios negreiros, a escravidão assumia um outro significado. Era como se esses distintos grupos passassem por um processo de duplo desenraizamento. O primeiro, quando eram tirados forçosamente de seus grupos de origem tendo que deixar suas regras sociais e tudo que marcava as suas identidades. A segunda, quando inseridos em ambientes novos, seja nos tumbeiros, seja nas novas terras, eram submetidos a novas regras sociais, a condições subumanas, de subordinação, de abuso dos corpos, de regras duras de convivência, exclusão e exploração.

Macedo (2020) descreve os africanos como aqueles que possuíam modelos sociais variados, desde sociedades que tinham modos de vida mais simples (chefaturas), a até outras sociedades consideradas mais complexas (reinos e impérios). Ainda segundo o autor, os africanos

[...] conheciam formas específicas de hierarquização social, com divisão social do trabalho, instituições próprias de governo, produção e circulação de riquezas na agricultura, pecuária e artesanato, além de sistemas de troca e de representação de valor nos negócios de médio ou longo alcance, centros administrativos e religiosos bem estruturados. (Macedo, 2020, p. 71).

Esse deslocamento forçado foi executado como uma tentativa, sem dúvidas, de apagar a história dessas pessoas que eram “livres” em suas terras de origem. Vale ressaltar que há também registros históricos que atestam para o fato de que, na própria África, mesmo antes da presença árabe já havia escravização<sup>2</sup>, no entanto, de ordem distinta, pois não tinha caráter de comércio e nem de tráfico. O processo de desenraizamento, geralmente, esteve e ainda está

---

<sup>2</sup> Antes, a escravidão existente na África era doméstica já que muitos dos que se tornaram cativos foram por causa das guerras. Os escravizados domésticos cultivavam terras e outros serviços e tinham direito a propriedade da terra, casar e ter filhos. A escravização também podia acontecer como pena para delitos ou em períodos adversos como a fome e seca.

atrelado às mudanças econômicas, sociais e/ou políticas. No caso dos africanos, na era da expansão colonial das grandes potências da época, a descoberta e o desejo de dominar novas áreas para exploração econômica colonialista geraram para esses povos destruição de seus antigos modos de vida e de produção, separaram e enfraqueceram as relações com o lugar de origem e passaram a ter histórias marcadas pela “exclusão” e dominação, com processos de violência dos mais cruéis. Não que na terra natal não existisse o desenraizamento, mas isso ficou bem mais evidente na história de privação com a transplantação para as Américas. Conforme Bosi (1987b, p. 23), “o fato é que as transformações históricas – de caráter social, econômico e cultural – não podem ser negadas. E o processo histórico afeta as vivências comunitárias, tidas como básicas para o enraizamento”.

Os cativos, como eram chamados, de alguma forma, foram forçados, ainda mesmo antes do embarque na África, a um processo de eliminação de aspectos que ligavam a sua ancestralidade, seus nomes, religiões e, também de uma memória coletiva. O que restava para esses povos traficados eram a lembrança e a esperança de serem homens livres novamente. Alencastro (2000, p. 144) afirma que o que houve com esses povos foi um processo de dessocialização:

o indivíduo é capturado e apartado de sua comunidade nativa, se completa com *despersonalização*, na qual o cativo é convertido em mercadoria na sequência da reificação, da coisificação, levada a efeito nas sociedades escravistas. Ambos os processos transformam o escravo em fator de produção polivalente, e apresentam-se como uma das constantes dos sistemas escravistas. (grifo do autor)

Nesse processo de aprisionamento e privação de liberdade, de acordo com Alencastro, mais de 10 milhões de africanos foram retirados das suas terras e 40% foram levados aos portos brasileiros como mercadorias de troca, “peças”. O Brasil foi o principal receptor de cativos dos países africanos de língua oficial portuguesa e os negros eram levados, em sua maioria, dos portos do litoral do Golfo da Guiné, de Luanda<sup>3</sup> e de Moçambique. Conforme Macedo (2020), em relação ao tráfico desenvolvido no oceano Atlântico,

nota-se o gradual fortalecimento de “lusobrasileiros”, isto é, portugueses estabelecidos no Brasil ou descendentes de portugueses nascidos no Brasil, no comércio dos portos de Luanda. Eram negociantes nascidos em Pernambuco, Bahia e, principalmente no Rio de Janeiro, que com o tempo passaram a disputar a atividade do tráfico negreiro. No mesmo instante, na África Central,

---

<sup>3</sup> Os portos de Luanda serviram de escaudouro de escravizados devido à sua posição estratégica (era uma baía e protegida por uma ilha) e despertou interesse dos holandeses, de portugueses e de seus descendentes que nasceram no Brasil.

surgiam grupos especializados em capturar e transportar os cativos para o litoral, “os pombeiros”. Isso prova que a regularidade do comércio escravagista promoveu sua crescente especialização, envolvendo diferentes grupos e atividades. (Macedo, 2020, p. 108).

Vale observar ainda que a importação dos escravizados coincidiu com a transplantação das línguas africanas já documentadas desde a primeira metade do século XVI chegando até o século XIX. Aqueles que foram trazidos ao Brasil eram oriundos de lugares diferentes da África com predominância de Angola e Moçambique e, desde os primeiros ciclos de transplantação, a Coroa Portuguesa já tinha a política de misturar diferentes etnias africanas com o objetivo de impedir que os escravizados se concentrassem numa mesma região, não importando se pertenciam a tribos inimigas. Antes de os negros embarcarem do porto em Angola, muitos deles passavam meses nos chamados “barracões” à espera de outros que eram trazidos do interior com os pombeiros até que a carga do navio completasse. Esse processo podia levar meses. Embora as línguas faladas pelos diversos grupos de africanos capturados para serem trazidos para a América tivessem proximidades tipológicas, muitas oriundas dos grupos banto, os cativos, para manterem contato entre si, muito provavelmente, utilizaram o quimbundo como a língua veicular para viabilizar a comunicação, conforme afirma Bonvini (2020).

Essa intensa concentração de etnias, atrelada ao plurilinguismo forçado, gerou consequências linguísticas bem antes desses povos aportarem no Brasil. Quanto a essas questões de imposição linguística inicial e de outras geradas ao longo da história de contato entre os povos, Lucchesi (2015, p. 22) assegura que a situação linguística polarizada do PB assenta-se diretamente com o processo social escravocrata:

Durante quase quatro séculos, um terço da população do Brasil, formada por senhores europeus, implementou um projeto colonial baseado na escravização da população autóctone e de povos africanos. Essa violenta divisão original da sociedade brasileira deu ensejo a um fosso linguístico entre a língua do colonizador e as centenas de línguas faladas por indígenas e africanos [...] os descendentes desses últimos foram abandonando a língua de seus antepassados e adotando como língua materna a língua do colonizador, conquanto guardassem as marcas de sua aquisição imperfeita e nativização mestiça. (Lucchesi, 2015, p. 22 - 23).

No que se refere às línguas africanas que foram trazidas ao Brasil, Bonvini (2020, p. 33) aponta que houve rupturas de ordem semântica e dialógica. Para ele, independentemente das identidades e origens, fica evidente a presença de línguas que foram arrancadas de seu nicho ecológico, já que do ponto de vista semântico, “não refletiam mais a realidade africana [...] não

tinham apoio na nova realidade”. Sobre o viés dialógico, “as línguas foram confrontadas com contatos linguísticos inabituais [...] sem esquecer a perda do seu valor identitário consecutivo ao nivelamento da língua resultante do desaparecimento das variantes dialetais”. Sobre essa situação, Couto (2021) afirma que

[...] praticamente tudo na língua pode ser encarado da perspectiva do contato. Contato de línguas é um tipo de interação, só que uma **interação** em que falante e ouvinte têm sistemas linguísticos diferentes. A interação é universal [...] Há **contato de línguas** quando representantes de determinado povo e respectiva língua se deslocam e se encontram com representantes de outro povo e respectiva língua. [...] o problema é que, no caso presente, não há uma língua comum que permita a intercompreensão. O que há é uma **interação interlinguística**, na qual se faz uma adaptação mútua ou unidirecional, dependendo das relações de força em presença (Couto, 2021, p. 179; grifo do autor)

Não se pode negar como o sistema de exploração colonial foi capaz de marcar fortemente as suas ex-colônias. No Brasil, por exemplo, apesar das sociedades terem sido construídas em um processo de diáspora africana, de um passado escravista, até hoje, elas tentam se conectar social e culturalmente, seja pelas diversas formas de manifestações artísticas ou não, pela ciência, pela culinária, pela religiosidade, pelo culto às divindades africanas, pela música, pelo resgate das línguas, pela autoafirmação da identidade negra. Essa africanidade é marcada, já há algum tempo quando, por exemplo, os africanos e afrodescendentes da diáspora, tentaram ressignificar os cultos e práticas religiosas no Brasil com a origem de “terreiros”, criaram irmandades como espaços que servissem de socialização e de construção de identidade para eles fora da África.

Considerando a forte ligação entre o Brasil e o continente africano, a expansão colonial portuguesa e o tráfico forçado de negros para o trabalho escravo, pretende-se, na subseção que segue, tecer algumas reflexões sobre as relações entre a língua e os povos. Também, são apresentadas algumas das consequências provocadas pelas expansões ultramarinas portuguesas para a sociedade colonial e pós-colonial, com enfoque nos efeitos linguísticos e, o consequente surgimento de variedades da língua portuguesa com peculiaridades, “originalmente”, brasileiras, fruto de violências e preconceitos de um colonialismo perverso. Para essa reflexão, outros olhares com base em estudos culturais são utilizados e o enfoque é dado partindo da sócio-história e de como os povos autóctones, suas línguas, suas culturas e dos africanos que aqui chegaram eram vistos.

### 2.2.1.2 Reflexões sobre a língua portuguesa em contato com os povos autóctones e africanos no Brasil colonial e independente

*As pessoas não associam necessariamente esses juízos com preconceito ou discriminação em termos de raça ou classe social: elas acreditam que, sejam quais forem as características sociais dos falantes, estes simplesmente usaram a língua de um modo errado e que existe para eles a possibilidade de aprender a falar corretamente. Se não o fizerem, é por culpa própria deles, como indivíduos, seja qual for sua raça, cor, credo ou classe; existe uma abundância de modelos do “bom” falar para eles. (MILROY, 2011[2001], p. 59)*

A língua é uma atividade, fundamentalmente, social e de elemento identitário, sendo assim, ela pode ser influenciada por vários processos socioculturais e históricos. Sobre isso, Couto (2021, p. 181) afirma:

[...] o deslocamento de povos e respectivas línguas para o território de outros povos e respectivas línguas provoca muitas outras situações **linguísticas complexas**. Uma das mais conhecidas é o **bilinguismo** e o **multilinguismo**, além é claro do **multidialealismo**. [...] São esses diversos tipos de contato que constituem a dinâmica das línguas. São eles que provocam a **evolução linguística**. [...] toda mudança na língua é causada por algum tipo de contato”  
[...] A **morte da língua (glototanásia)** é uma das consequências mais lamentáveis do contato de línguas nos tempos atuais. (grifo do autor)

É necessário examinar os diferentes contextos sócio-históricos para a formação da língua portuguesa em cada um de seus territórios, com suas diferenças, embora os laços estruturais existam. Segundo Mattos e Silva (2008, p. 4),

[...] quando se fala de língua portuguesa una, na diversidade de suas manifestações, tem-se em mente - por um lado - a estrutura comum que está na base das suas diversificadas realizações e por outros os factores históricos que a unem e a definem como tal.

Quando os portugueses chegaram ao Brasil, o país já era povoado pelos povos originários e só depois, uma grande quantidade de africanos escravizados foi trazida pelos portugueses. Por muito tempo, o português europeu, os indígenas e os negros africanos constituíam a base da população brasileira. A princípio, a colonização se deu no litoral. Com a fundação de São Paulo, foi aberta uma porta de entrada para o processo de interiorização do país. Nesse contexto, o Brasil, já no século XIX, inicia um processo de urbanização, escolarização e industrialização após a chegada da corte lusitana.

De forma mais acentuada em relação com o que aconteceu com o português europeu, nas outras regiões para onde a língua portuguesa foi levada, o português esteve inserido em situações complexas e diversas de multilinguismo e, desde os primeiros momentos de colonização, a situação sociolinguística no Brasil diferiu daquela vivida pelos angolanos, por exemplo. Enquanto nas terras brasileiras os portugueses assumiram desde cedo o seu papel de colonizador e exploraram as terras assim que chegaram, em Angola, os portugueses eram, a princípio, “parceiros” comerciais. Isso posto, antes mesmo de iniciar a transplantação dos africanos para o continente americano, os portugueses já mantinham relações comerciais com populações do continente africano e, só um pouco mais tarde, países africanos também se tornariam colônias do império português.

Há de se considerar, por exemplo, que mesmo que Brasil e Angola tenham peculiaridades em suas trajetórias sócio-históricas, segundo Teixeira (2017, p. 15) esses dois países têm suas histórias de colonização entrelaçadas e, desde o século XVI, também já existia a presença de brasileiros em território angolano. A autora afirma:

Angola foi a *mina de ouro* dos portugueses, com a qual abasteceram as plantações de cana de açúcar e os engenhos de mão de obra escrava, sobretudo na Bahia, nos primeiros séculos da colonização. Se por um lado, era mais oneroso buscar essa mão de obra na África, por outro, os africanos poderiam se mostrar mais produtivos na medida em que já estavam acostumados ao trabalho de eito. Além disso, viriam para as terras desconhecidas, o que era outra vantagem sobre os naturais. Adicionalmente, com a facilidade de navegação costa à costa, Brasil/África/Brasil o comércio de almas resultou em uma fonte de riqueza para luso-brasileiros destemidos e gananciosos. Era para lá que também eram encaminhados os degredados do Brasil, como foi o caso do poeta baiano Gregório de Matos, desterrado em Benguela. (Teixeira, 2017, p. 15)

No Brasil, inicialmente, não houve uma oposição dos nativos à colonização como houve em Angola. Há um contexto rápido de aculturação dos povos autóctones e, em termos linguísticos, houve a predominância do uso das línguas gerais de base tupi guarani pelos portugueses, indígenas e seus descendentes, a ponto de, apenas no século XVIII, o Marquês de Pombal, em 1759, expulsar os jesuítas e instituir o português como língua oficial da Colônia, dentre outras medidas determinadas pelo Decreto Pombalino. A expulsão dos jesuítas foi o que lhes permitiu um controle em termos de políticas linguísticas no Brasil.

Nesse período, linguisticamente falando, Teyssier (1997, p. 94) descreve a fala brasileira da seguinte forma:

Os “colonos” de origem portuguesa falam o português europeu, mas evidentemente com traços específicos que se acentuam no decorrer do tempo.

As populações de origem indígena, africana ou mestiça aprendem o português, mas manejam-no de uma forma imperfeita. Ao lado do português existe a *língua geral*, que é o tupi, principal língua indígena das regiões costeiras, mas um tupi simplificado, gramaticalizado pelos jesuítas e, destarte, tornado uma língua comum. (Teyssier, 1997)

Conforme Cunha (1986, p. 202-203), as variedades portuguesas faladas que vieram ao Brasil de diferentes regiões de Portugal possuíam matizes mais conservadoras ou inovadoras a depender das regiões de onde provinham. Ainda, segundo o autor, o desenvolvimento da língua portuguesa na “América Lusitana<sup>4</sup>” tinha condições mais propícias à conservação do que a renovação de suas formas da língua, pois viveu mais de trezentos anos sem contato com outros povos, sem imprensa, sem núcleos culturais de importância, com poucas escolas e só, no século XVIII, esse panorama de conservadorismo intenso<sup>5</sup> foi alterado com o *rush* para as Minas.

Desde o início da colonização, o português foi aprendido em uma situação de aquisição imperfeita do português como segunda língua (L2). Somando-se a isso, havia uma ausência, uma precarização do sistema educacional no Brasil colonial. Silva Neto (1950, p. 36), com uma concepção impregnada de preconceitos, apresenta os povos indígenas e negros como povos que não atingiram um grau de civilização, ou seja, “em atrasado estágio de civilização”, comparado como o que o português europeu tinha atingido, não eram urbanos, nem civilizados porque não possuíam literatura. O falar dos indígenas e europeus, para Silva Neto (1950, p. 96) eram [...] “cicatrizes da tosca aprendizagem que da língua portuguesa, por causa da sua mísera condição social”

Essa divisão presente na sociedade brasileira colonial deixa um hiato linguístico entre a língua do colonizador e das outras línguas faladas pelos indígenas e pelos africanos escravizados. Devido ao processo de aculturação, tanto os indígenas quanto os negros escravizados foram forçados à submissão linguística e seus descendentes aos poucos foram abandonando as suas línguas originárias e passaram a adotar a língua do colonizador. Desse modo, não houve, para esses povos, apenas a violência física, mas também uma violência simbólica e cultural. É possível interpretar a oposição à supremacia do português europeu como uma forma de autonomia cultural e uma defesa das tradições linguísticas dos antepassados. Esta abordagem é frequentemente impulsionada por uma compreensão crescente do valor da

---

<sup>4</sup> Termo utilizado pelo autor e que não reflete a posição tomada nesse trabalho.

<sup>5</sup> Logicamente, o posicionamento do autor deve ser relativizado, pois, se, de fato, a intensificação do contato entre línguas ocorreu a partir das descobertas de ouro e pedras preciosas em Minas Gerais, não se pode ignorar as inúmeras situações de contatos linguísticos que já havia no Brasil nos séculos anteriores.

diversidade linguística na melhoria do legado cultural e na promoção de uma identidade nacional que inclui o passado colonial e africano.

Depois da vinda da família real portuguesa, houve a mudança na escolarização e um posterior processo de urbanização das cidades favoreceu algumas mudanças linguísticas visto que falantes da norma popular levaram essa variante para as áreas urbanas. Segundo Mattos e Silva (2004, p. 20),

uma série de fatores de história externa conduzem à definição do Brasil como país majoritariamente de língua nem indígena nem africana. O multilingüismo menos ou mais generalizado, a depender da conjuntura histórica local nos séculos anteriores, localiza-se e abre, então, o seu caminho o português brasileiro (Mattos e Silva, 2004)

O processo de urbanização e industrialização, no Brasil, mudou um cenário que antes era, essencialmente, rural. Conforme Cunha (1986, p. 203), as cidades se tornaram centros criadores e difusores da cultura e mudaram com sua força normalizadora o tabuleiro linguístico do Brasil. Assim, “[...] apesar de viver uma fase essencialmente inovadora, [...] a variante brasileira [...] continua a apresentar características que, tendo sido também da variante portuguesa, no século XVI ou, ainda, no século XVII, posteriormente nela se perderam”. Conforme se verá adiante, o português nas áreas mais rurais de Angola, a linguagem dos povos autóctones das línguas nacionais se comparado às áreas rurais brasileiras, dos indígenas, dos mestiços e dos negros também sofrem mais resistência em relação à língua oficial, dita culta pelos portugueses das grandes cidades.

Para corroborar com essa discussão, Lévi-Strauss (2008) reitera o fato de os homens não aceitarem as diversas culturas como um fenômeno natural e frutos das relações entre os povos de forma direta ou indiretamente. Seu pensamento, diferentemente do que achava sobre os povos autóctones e dos negros africanos no Brasil no período colonial, é contrário a qualquer ideia de superioridade que se tome a civilização ocidental. Ele critica a atitude dos que se dizem “civilizados” se comportarem exatamente como os “bárbaros” ou “selvagens”. A considerada superioridade cultural portuguesa era vista à frente das outras culturas e, também, era atrelada à valorização da modalidade escrita da língua como marca de unidade e preservação do patrimônio cultural português, em detrimento da falada, considerada inferior e deturpada na visão dos portugueses.

Sobre a visão que os colonizadores tinham quando se depararam com a população de indígenas e, em seguida de negros trazidos da África, Silva Neto (1976) trata a inclusão dos indígenas e dos negros, que forçados chegaram para o trabalho escravo na sociedade como um

dos problemas, cujas línguas, segundo ele, revelavam uma linguagem rude de gente inculta, chamada de *crioulo* ou *semicrioulo*. A orientação ideológica presente na obra de Serafim da Silva Neto parte do ponto de que existe uma língua de uma cultura que é superior, a portuguesa, em detrimento das línguas autóctones e africanas.

Silva (2012) afirma que tanto a identidade quanto a diferença são relações sociais e, por serem discursiva e linguística, estão sujeitas a relações de poder que são impostas e nunca inocentes. Assim, as marcas da presença do poder podem ser traduzidas nas seguintes relações: inclusão/exclusão (os que pertencem e não pertencem); de demarcar fronteiras como (nós *versus* eles); classificação (bons e maus; puros e impuros; desenvolvidos e primitivos; racionais e irracionais; normais e anormais. É nítido, então, essa diferenciação na relação entre os colonizadores e os colonizados na história sociolinguística do Brasil. Desse modo, no processo de polarização, ou seja, de relações binárias, vê-se que um dos termos é sempre o privilegiado, enquanto o outro é visto de maneira inferior e com uma carga negativa.

Os preconceitos racial, social e econômico estão bastante entranhados na sociedade da época e, segundo Frans Boas (2010), um antropólogo teuto-americano, o preconceito racial deve ser contestado. Afirma que a mistura entre grupos humanos distintos não significa degeneração e muito menos risco às raças ditas superiores, mas é, sobretudo, um desempenho muito importante para a história das populações modernas. Conforme Boas (2010) não existem diferenças raciais, mas sim, diferenças culturais, étnicas. Cada cultura, ou seja, cada grupo humano possui determinado caráter que se expressa no seu comportamento, sendo que os diferentes comportamentos não podem ser atribuídos aos fatores genéticos, mas às condições sociais, ao ambiente cultural. Além do mais, Boas (2010) discorre sobre as populações modernas com o conceito de “relativismo cultural” em que a cultura é um termo plural e não singular já que é impossível encontrar leis gerais, universais da “evolução” das culturas. Não há, portanto, diferença na natureza biológica dos primitivos e civilizados, assim, não são inatas, mas adquiridas culturalmente. Busca-se, então, tirar o estudo de uma cultura particular, o etnocentrismo (julgamento de valor) e aceitar cada cultura com seu “estilo” próprio, particular.

No período que compreende o Brasil independente, a partir de 1822, o país passa a ser influenciado também por outros povos, como os alemães, italianos e franceses. Angariar proletários de outras partes do mundo era a alternativa encontrada para substituir os escravizados mortos, os fugidos e aqueles que deixaram de vir da África. Esse novo contexto possibilitou a vinda maciça de imigrantes europeus e contribuía para tal “branqueamento” do Brasil. O papel dos imigrantes europeus foi algo estratégico para a sociedade que estava em formação. Segundo Croci (2011, p. 83),

[...] na receita para a formação da nação o ingrediente branco era tolerado exatamente enquanto componente indispensável para “melhorar a raça”, mas sempre em condição subalterna. Os imigrantes, brancos europeus, portanto de raça superior, podiam ser integrados no projeto identitário da elite nacional somente se *funcionais e subalternos* a ele. (Crocí, 2011)

Com uma população brasileira composta de mais de 2/3 por populações não-brancas, mais uma vez, fica evidente a luta de interesses e a imigração se torna um instrumento de “civilização”. Assim sendo, os colonos recém-chegados não serviriam apenas como mão de obra, mas também como um agente revigorante e de “melhoramento da raça”. Vale ressaltar que, atualmente, a situação da população negra no Brasil é muito preocupante, pois ela ainda é acometida pelo racismo e pelo preconceito estruturais. Não é porque houve a abolição da escravidão que os negros se libertaram dessa discriminação histórica e cultural nas terras brasileiras. Estudos mostram que a população negra tem sofrido com a desigualdade social e isso é refletido em salários mais baixos, maiores números de desemprego no país, maiores taxas de analfabetismo e maior população carcerária do país, só para exemplificar algumas injustiças. Há ainda casos em que muitos são flagrados vivendo em situações análogas à escravidão.

Há de se levar em conta, portanto, que a dispersão de milhares de negros pelo território do Brasil, falantes de muitas línguas e grupos étnicos diferentes, foi decisiva para o desenvolvimento do português brasileiro atual, visto que a população mestiça e negra, após longos três séculos de contato, representavam ainda na época da Independência do país, 75%. Mattos e Silva (2016) afirma que o português brasileiro origina-se, em linhas gerais, do europeu no Brasil, mas se caracterizava de uma forma na “complexa interação entre a língua do colonizador (e, portanto, do poder e do prestígio), as numerosas línguas indígenas brasileiras, as numerosas línguas africanas chegadas pelo tráfico negreiro”, bem como, “as línguas dos que emigraram para o Brasil da Europa e da Ásia, sobretudo, a partir de meados do século XIX”.

Contextualizando sobre as realidades sociolinguísticas brasileira e angolana, é necessário considerar que Brasil e Angola, por exemplo, vivenciaram uma espécie de “rompimento” das suas identidades. As línguas foram destroçadas entres os grilhões e o despejo social. A língua portuguesa das colônias apresenta marcas gêmeas cunhadas pelo colonialismo, cicatrizes que ainda latejam nas palavras dos excluídos. A história mostra que a força desenraiza e despersonifica, justificando assim a dominação. A imposição da vontade de manutenção do *status quo* cala aquele que é subjugado e ecoa nas gerações os gritos das desigualdades que não refletem a diversidade cultural e linguística de povos e nem respeitam as suas origens. Entender

tais processos é buscar uma possível reconexão ao passado e perceber hoje como a construção identitária é dinâmica.

Na subseção seguinte, será discutido sobre quem eram os africanos que foram trazidos dos tumbeiros nos principais portos brasileiros, suas procedências, as rotas utilizadas, os grupos que chegaram e as línguas que trouxeram. Não somente, é mostrada a distribuição populacional brasileira ao longo dos 300 anos, mas também a distribuição geográfica dos negros escravizados relacionando-os com as atividades por eles exercidas. A subseção também destaca o quão intenso e significativo foi o contato dos negros com os indígenas e portugueses, em sua maioria, oriundos de grupos bantu na emergência de uma caracterização de novas variedades da Língua Portuguesa no Brasil, em especial, das variedades mais populares, o PVB (doravante Português Vernacular Brasileiro).

### 2.2.2 As línguas bantu e a emergência de uma variedade brasileira do português

Das línguas que foram transplantadas para o Brasil, Bonvini (2020) afirma que os povos africanos procediam, basicamente, de duas áreas. A primeira, a área oeste-africana, que tinha um maior número de línguas, no entanto, muito diferenciadas entre elas. A segunda, a área austral, que era, essencialmente, do subgrupo banto. A princípio, essa área limitava-se mais à costa oeste do continente - República Democrática do Congo e Angola, e, só mais tarde, abarca também a costa leste (Moçambique). Na área austral, as línguas tinham características bem diferentes da primeira área pois, embora fossem bastante reduzidas em termos quantitativos, eram faladas por uma quantidade maior dos negros capturados para o serviço escravo, os cativos. O Quadro 1, a seguir, mostra um panorama das áreas de origem dos cativos:

**Quadro 1: Áreas de procedências dos cativos**

Área Oeste-africana	Área Leste-africana
<p><b>atlântica:</b> <i>fula (fulfulde), uolofe, manjaco, balanta;</i></p> <p><b>mandê</b> (sobretudo, o <i>mandinga</i>); <i>bambara, maninca, diúla;</i></p> <p><b>gur:</b> <i>subfamilia gurúnsi;</i></p> <p><b>cuá</b> (subgrupo <i>gbe</i>): <i>eve, fon, gen, aja</i> (designadas pelo termo <i>jeje</i> no Brasil);</p> <p><b>ijóide:</b> <i>ijó;</i></p>	<p>H.10 <i>congo (quicongo): quissolongo, quissicongo (quissangala), quizombo, quissundi</i> (falada pelos bacongos, numa zona correspondente ao antigo reino do Congo) e <i>quivili, iuoio (fiote), quiombe</i> (faladas em Cabinda e em Loango);</p> <p>H.20 <i>quimbundo</i> (falada pelos ambundos, na região central de Angola, correspondendo ao antigo reino de Ndongo), <i>quissama, quindongo;</i></p>

<p><b>benuê-congolesa:</b></p> <p><b>defóide:</b> falares <i>iorubás</i> designados no Brasil pelo termo <i>nagô-queto</i>;</p> <p><b>edóide:</b> <i>edo</i>; <i>nupóide:</i> <i>nupe (tapa)</i>;</p> <p><b>ibóide:</b> <i>ibo</i>;</p> <p><b>cross-river:</b> <i>efique, ibíbio</i>;</p> <p><b>afro-asiático:</b> <b>chádica:</b> <i>hauçá</i>;</p> <p><b>nilo-saariano:</b> <b>saariana:</b> <i>canúri</i>.</p>	<p>H.30 <i>iaca-holo:</i> <i>iaca, imbangala, chinji</i>;</p> <p>K.10 <i>chôcue:</i> <i>uchôcue, ochinganguela, chilucazi, luena (luvale)</i>;</p> <p>L.30 <i>luba:</i> <i>chiluba-cassai (lulua)</i>;</p> <p>L.50 <i>lunda:</i> <i>chilunda, urunda</i>;</p> <p>P.30 <i>macua:</i> <i>omacua</i>;</p> <p>R.10 <i>umbundo</i> (falado pelos ovimbundos na região de Benguela, em Angola): <i>umbundo, olunianeca</i>;</p> <p>R.20 <i>cuaniama, indonga:</i> <i>ochicuaniamama, cuambi</i>;</p> <p>R.30 <i>herero:</i> <i>ochiherero</i></p>
-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Fonte: adaptado de Bonvini (2020, p. 30-31)

Dos mais de quatro milhões de pessoas que aportaram no Brasil no período escravocrata, a maioria dos chegados, cerca de 75%, eram falantes das línguas provenientes da área oeste, como as áreas dos reinos do Congo e Ndongo e, que falavam o quimbundo e o quicongo. No início do tráfico negreiro, segundo Castro (2011, p. 4) os cativos falantes dessas duas línguas podiam ser encontrados em várias partes do território brasileiro no século XVI, e, a partir do século XVIII, se juntam aos primeiros os falantes do umbundo, da região de Benguela (Angola), em quantidade relativamente menor, mas com igual importância.

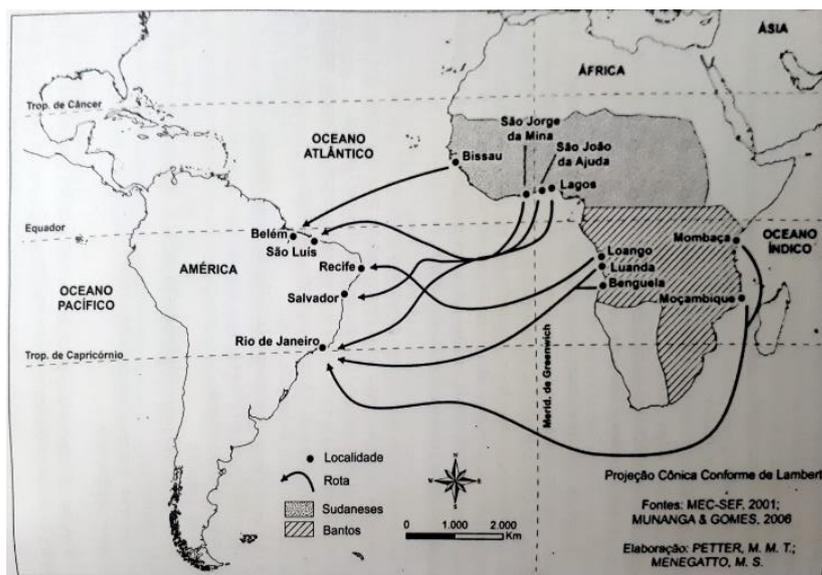
Um dos fatores que fez com que os indivíduos da área mais oeste fossem mais presentes no mercado transatlântico se explica pela maior facilidade de reunir um maior número de cativos. Por ser uma região mais costeira, era privilegiada, já que também os cativos que vinham de áreas mais interioranas representavam mais riscos (sentiam-se mais fracos da viagem do interior para a área litorânea e dava mais trabalho para os pombeiros).

Na Figura 1, são mostradas as rotas utilizadas pelos cativos para chegarem em solo brasileiro. Conforme se observa, os navios negreiros que saíam do território africano partiam principalmente, de quatro rotas, a saber: Rota de Moçambique, Rota de Angola, Rota da Mina e Rota da Guiné. As embarcações utilizadas eram superlotadas e, em média, 300 a 600 cativos eram aprisionados nos tumbeiros, a depender do tipo de embarcação. Os africanos eram caçados pelos pombeiros, que recrutavam os cativos em trocas de mercadorias como tecidos, armas de fogo, tabaco, ferro, aguardente, marfim entre outros produtos. As travessias eram longas e duravam meses para chegar ao destino, as condições sanitárias enfrentadas pelos aprisionados

eram precárias e estima-se que 20% morriam devido ao acometimento de doenças e pelo ambiente insalubre.

A Rota de Angola trouxe aproximadamente quase metade dos africanos que chegaram na América, que partiam dos atuais territórios do Congo e de Angola para os portos de Recife, Salvador e Rio de Janeiro e começou ainda no século XVI. O tráfico oriundo da Rota da Guiné, iniciou também no século XVI cujos escravos aportavam nas regiões mais ao norte no território brasileiro (Norte e Nordeste). Essa rota foi considerada a que menos teve um impacto na formação da população brasileira se comparada às outras rotas, pois o movimento ainda era bastante pequeno quando ela foi iniciada. Na Rota da Mina, Mina se tornou um entreposto importante de escravos no século XVII que abasteceu os portos brasileiros nos estados do Maranhão, Rio de Janeiro e, principalmente, a cidade de Salvador (Bahia) durante uns 200 anos. Já a Rota de Moçambique, uma rota que foi pouco explorada por algum tempo devido às longas distâncias, começou suas atividades no início do século XIX quando Portugal precisou fugir da pressão da Inglaterra pelo domínio do Atlântico. Os cativos dessa rota saíam da capital Maputo (Porto de Lourenço Marques) e se dirigiam para o Rio de Janeiro.

**Figura 1: Principais rotas do tráfico transatlântico**



Fonte: Adaptado de (Munanga, 2009 *apud* Petter, 2015, p. 224)

A área bantu compreende falantes de mais de 300 línguas e, são faladas em 21 países do continente africano. Por toda sua história, pela sua antiguidade, pelo grande número de falantes e pela sua grande distribuição em terras brasileiras durante o período colonial, percebe-se que os povos do grupo linguístico banto foram um dos maiores responsáveis pela formação do PB.

O povo africano é, assim, considerado por muitos autores, a exemplo de Castro (2001), Mattos e Silva (2004), o principal agente transformador da variedade brasileira, com destaque ao português popular falado, pois, muitos daqueles que chegaram aqui adquiriram o português como segunda língua e as interferências de suas línguas maternas fizeram surgir uma variedade que se distingue do PPE (Português Padrão Europeu).

Com o intuito de auxiliar na evangelização e doutrinação, os jesuítas precisavam entender e aprender a língua dos nativos, a língua geral. Deste modo, os jesuítas acabaram escrevendo diversas obras, dentre algumas delas, *A Arte da língua de Angola*, escrita pelo padre Pedro Dias, representava uma prova histórica de que uma língua africana era empregada no Brasil. Esse documento foi escrito em 1697 com redação no Brasil e publicação em Lisboa. Segundo Bonvini (2020), além desse documento permitir a descoberta de uma política linguística utilizada pelos jesuítas, tinha, sobretudo, não só uma relevância linguística, como também aponta ainda três pontos de vista sobre a obra:

[...] é objetivamente a primeira descrição do quimbundo, redigida, além disso, a partir de dados recolhidos no próprio Brasil; contém, em seu texto, observações que permitem mostrar o olhar que um falante do português dessa época lançava sobre uma língua africana tipologicamente diferente da sua; é enfim o testemunho do português utilizado no século XVII no Brasil, nos meios cultos. (Bonvini, 2020, p. 34)

*A Arte da língua de Angola*<sup>6</sup>, a gramática mais antiga de uma língua banto, a primeira gramática sistemática do quimbundo, já fazia referência a uma língua cuja fronteira já ultrapassava o território da Bahia. Com isso, muitos negros que foram escravizados e que estavam dispersos no solo brasileiro já utilizavam essa língua como uma língua veicular, um tipo de língua franca que era utilizada para facilitar a comunicação por pessoas oriundas de diferentes etnias, culturas e línguas, já empregada em Luanda.

É interessante ressaltar que, mesmo com o surgimento de uma gramática do quimbundo, os cativos não, necessariamente, falavam apenas o quimbundo visto que os embarques que aconteciam em solo africano traziam negros de diferentes partes de Angola. Assim, Bonvini (2020) destaca que os cativos que vieram de áreas bantas presenciaram uma experiência linguística bastante diferenciada, estavam sujeitos a um plurilinguismo forçado e concentracionário, já que eles, ainda em seus barracões eram, propositalmente, “misturados”,

---

<sup>6</sup> A gramática tinha as seguintes informações de catalogação, conforme Bonvini (2020, p. 33), “*Arte da lingua de Angola, oeferecida a Virgem Senhora N. do Rosario, Mãe, e Senhora dos mesmos Pretos, pelo P. Pedro Dias da Companhia de Jesu*”. Lisboa, na Officina de Miguel Deslandes, Impressor de Sua Magestade. Com todas as licenças necessarias. Anno 1697. 48 p.”

antes mesmo de embarcarem para o Brasil, e já utilizam o quimbundo como uma língua veicular nessas concentrações forçadas e prolongadas entre os falantes de línguas africanas diferenciadas mas que eram, tipologicamente, próximas. Ao chegarem ao Brasil, o cenário não foi muito modificado.

Como já citado anteriormente, o tráfico negreiro trouxe uma grande quantidade de africanos ao Brasil. Conforme se vê no Quadro 2, a seguir, elaborado por Alberto Mussa partindo de fontes sobre a diacronia da população brasileira, em todo o período entre o século XVI e o século XIX. Segundo o autor, a população brasileira era composta, em sua maioria, pelos negros de origem banta, sendo que, entre os séculos XVII e XVIII, a presença desses povos teve um aumento considerável chegando a mais de 60% da população brasileira na época, permanecendo “estável” por quase três séculos.

**Quadro 2: Línguas africanas trazidas pelo tráfico ao Brasil (Século XVI ao XIX)**

	XVI	XVII	XVIII	XIX
Oeste-atlântico	20%	7%	1%	3%
Mande	20%	7%	1%	3%
Kru	14%	5%	1%	3%
Gur	1%	3%	8%	9%
benue-kwa (não banto)	7%	10%	<b>20%</b>	<b>24%</b>
Banto	<b>35%</b>	<b>65%</b>	<b>64%</b>	<b>50%</b>
outros (adamawa-ubanguiano; dogon: não nigercongo)	3%	3%	5%	8%

Fonte: Mussa (1991, p. 145 *apud* Mattos e Silva, 2004, p. 18) (destaque nosso)

Ainda, quanto à questão das línguas africanas em terras brasileiras, Mussa afirma que, embora os bantos estivessem mais presentes, os não bantos, mesmo de uma forma diferente, também contribuíram para a formação do PB, principalmente nos dois últimos séculos com 20% e 24%, respectivamente. Assim,

[...] o percentual de falantes Bantu foi sempre superior, e quase sempre maciçamente, em todo o período do tráfico. Isso nos possibilita entender de forma bastante clara por que são precisamente os itens lexicais de origem banto os que se registram com mais anterioridade, com maior grau de integração morfológica e em maior número de campos semânticos no português do Brasil... a posição relativamente proeminente do benue-kwa (não banto) nos últimos séculos também implica o grande número de itens lexicais emprestados por essas línguas, embora não integrados e particularmente

restritos aos campos semânticos ligados à atividade ritual. (Mussa, 1991, p. 146 *apud* Mattos e Silva, 2004, p. 18).

Em um outro quadro do mesmo autor, o Quadro 3, a seguir, Mussa traz um retrato de como a população brasileira era constituída desde o século de povoamento na costa brasileira até o período que acontece a abolição do tráfico negreiro. Inicialmente, a população é composta, em grande parte, pelos índios integrados que conviviam com os portugueses recém-chegados, muitos deles escravizados e não falantes da língua portuguesa, enquanto a população negra ainda é minoria. Com a intensificação do tráfico de escravos oriundos da África, a população brasileira vai ganhando contornos diferentes, visto que, já no século seguinte, era composta por 50% de africanos e negros brasileiros que já falavam o português de uma forma “modificada”, em decorrência de terem adquirido o Português como segunda língua.

**Quadro 3: Distribuição demográfica no Brasil do século XVI-XIX**

	1538-1600	1601-1700	1701-1800	1801-1850	1851-1890
Africanos	20%	30%	20%	12%	2%
negros brasileiros	-	20%	21%	19%	13%
mulatos	-	10%	19%	34%	42%
brancos brasileiros	-	5%	10%	17%	24%
européus	30%	25%	22%	14%	17%
índios integrados	50%	10%	8%	4%	2%

Fonte: Mussa, 1991, p. 163 *apud* Mattos e Silva, 2004, p. 18)

Nesse cenário, percebe-se uma direção inversa da demografia inicial. À medida que os séculos vão passando, há o crescimento de mulatos e dos brancos brasileiros. A miscigenação entre os personagens dessa história vai se intensificando, e, conseqüentemente, há uma quantidade maior no número de mulatos; a população indígena é, em grande parte, dizimada e decresce vertiginosamente de 50% para 2%; a presença cada vez menor de portugueses já são fatores que ajudam a delinear uma variedade legitimamente cada vez mais brasileira da língua portuguesa.

Pensando em termos quantitativos e na distribuição dos cativos pelas cinco regiões brasileiras, em 1819, as regiões que mais concentravam escravos eram, de acordo com Mattoso, (1990, p. 65 *apud* Mattos e Silva 2004, p. 104), a região Centro-Oeste (40,7%) seguida pelas regiões Nordeste (33%) e Sul (28,9%). Não tão diferente, a região Leste tinha 28,1% e, em menor quantidade, a área Norte do país com 27,3%. Devido a uma nova dinâmica social,

observa-se que esses dados apontam para uma distribuição mais equilibrada na população de escravizados do Brasil no final da colonização.

Nesse aspecto, já é possível notar esse novo cenário da distribuição dos escravizados no Brasil, já em vias da supressão do tráfico africano em 1830, forçou os escravizados a migrarem para outras regiões, inclusive, para áreas mais urbanas. Onde antes havia uma maior concentração de escravizados nas lavouras de cana-de-açúcar em Pernambuco, Bahia e Rio de Janeiro (nos séculos XVI e XVII), na mineração de ouro e diamantes em Minas Gerais, Mato Grosso e Goiás (final do século XVII), a atração da mão-de-obra migrou-se para o litoral do país, Rio de Janeiro e São Paulo (século XIX), sobretudo, por causa da atividade cafeeira e alguns outros cultivos pelo território do Brasil com o algodão, o fumo, entre outros.

No Quadro 4, adiante, Mattos e Silva (2004) mostra números populacionais de escravizados de 1884, com base em Conrad (1978). Esses dados corroboram o cenário que já se verificava em 1819 às vésperas da abolição da escravatura que continuou seguindo na mesma direção, com a população escrava não mais na região Centro-Oeste e Nordeste, se comparado com o fim do período colonial.

**Quadro 4: População escravizada**

<b>Províncias</b>	<b>1884 / população escrava</b>
Minas Gerais	310.000
Rio de Janeiro	260.000
São Paulo	160.000
Bahia	140.000
Pernambuco	80.000
Rio Grande do Sul	60.000
Maranhão	45.000

Fonte: Conrad (1978, p. 81 *apud* Mattos e Silva, 2004, p.105).

A figura 2 mostra um mapa de Pessoa de Castro (2006, p. 113) para a etnologia africana no Brasil que se entrelaça com o que já está sendo descrito aqui. No mapa, é possível ver a distribuição dos grupos africanos pelos estados brasileiros, conforme se atenta à legenda na parte inferior esquerda dele. Desse modo, assim que a introdução maciça de escravizados africanos chega ao Brasil para trabalhar, tem-se, exclusivamente, indivíduos oriundos de grupos bantus e, se localizam, basicamente, por todo o território do Brasil, no campo da agricultura.

Com o passar do tempo, escravizados de outros grupos linguísticos vão também ocupando espaço, como os jeje-mina, o nagô-iorubá e os hauçá (estes apenas no Século XIX), sobretudo, a partir do século XVII, e dividem espaço com os bantos, estando correlacionados com diferentes tipos de atividades econômicas exercidas, conforme a Figura 2, a seguir:

**Figura 2: Esboço de mapa etnológico africano no Brasil**



atividade principal	século de introdução maciça			
	XVI	XVII	XVIII	XIX
agricultura	B	B/J	B/J/N	B/J/N
mineração			B/J	
serviços urbanos				B/J/N/H

Fonte: Pessoa de Castro (2006, p. 113)

À medida que as atividades desempenhadas vão se diversificando com o fim do império português no Brasil colonial, já se percebe, no século XIX, um quadro em que todos os grupos bantos apontados por Pessoa de Castro (2006) continuaram a trabalhar na agricultura e nos centros urbanos, com a presença dos hauçá. Percebe-se também que a mineração deixou de ser uma atividade atrativa para esses povos. Embora a colônia ainda tivesse pouca urbanização, os escravizados que trabalhavam desempenhando serviços urbanos. “Em 1820, apenas 7% da população vive nas cidades” (Mattoso, 1990, p. 109 *apud* Mattos e Silva, 2004, p.104). Nos centros urbanos, os escravizados passaram a desempenhar as funções de artesãos, carregadores, pintores e marinheiros. Na época, ainda existiam aqueles que trabalhavam em serviços domésticos ou vendiam produtos em tabuleiros. Mattos e Silva (2004, p. 106) conclui que:

A presença maciça dos africanos e afro-descendentes que a demografia histórica demonstra; a atuação constante dos escravos nas grandes frentes de

economia da colonização; a mobilidade geográfica, decorrente das vicissitudes vida econômica de seus senhores e da economia brasileira; os diversificados e múltiplos papéis desempenhados na sociedade colônia rural e urbana; o significado social e linguístico dos espaços ilegítimos da escravidão permitem embasar o meu ponto de vista interpretativo de que é esse segmento numeroso e operante – os africanos e afro-descendentes – o agente principal da difusão do português no território brasileiro, na sua majoritária, a popular ou vernácula. (Mattos e Silva, 2004, p. 106).

Vale reforçar aqui que, nesse mesmo período, o Brasil recebe a chegada de outros 15.000 portugueses para contribuírem na aceleração do processo de “relusitanização” no Rio de Janeiro. Assim, o período que compreende o Brasil independente, a partir de 1822, já se caracteriza como um país influenciado também por outros povos, como os alemães, italianos e franceses através de sua cultura. A vinda maciça de imigrantes europeus contribuiu ao “branqueamento” do Brasil. Nesse ínterim, com já foi descrito, houve cessamento do tráfico de negros africanos e o aniquilamento dos indígenas no que se considera uma mestiçagem brasileira.

Não se deve perder de vista que a história do negro no Brasil começou muito antes dele ser trazido à América para ser cativo. Por três longos séculos, os africanos foram colocados em situações sub-humanas, numa viagem transatlântica que muitos mal conseguiam cumprir a jornada de viagem para se submeterem, forçadamente, aos caprichos e anseios de navegadores europeus, os exploradores portugueses. Na seção 3 desta tese, será possível conhecer um pouco mais da história dos povos autóctones em solo africano antes de serem trazidos para o Brasil e como as suas trajetórias foram se delineando.

Voltando a mencionar um pouco mais dos aspectos linguísticos africanos no território brasileiro, entre os períodos de 1550 e 1700, já era possível perceber, na língua portuguesa, diferenças em relação ao Português Europeu com seus empréstimos lexicais e, principalmente, a variação na estrutura do sistema fonético-fonológico do que era falado no português. Conforme Pessoa de Castro, no seu trabalho intitulado *Influência das línguas africanas no português brasileiro*<sup>7</sup>,

Essa interação lingüística, apoiada por fatores favoráveis de ordem sociohistórica e cultural, foi provavelmente facilitada pela proximidade relativa da estrutura lingüística do português europeu antigo e regional com as línguas negro-africanas que o mestiçaram. Entre essas semelhanças, o sistema de sete vogais orais (a, e, ê, i, o, ô, u) e a estrutura silábica ideal (CV.CV)

---

<sup>7</sup> Texto “Influência das línguas africanas no português brasileiro”, artigo encontrado no site: <http://smec.salvador.ba.gov.br/documentos/linguas-africanas.pdf>

(consoante vogal.consoante vogal), onde se observa a conservação do centro vocálico de cada sílaba, mesmo átona. Esse tipo de aproximação casual, mas notável, provavelmente possibilitou a continuidade do tipo prosódico de base vocálica do português antigo na modalidade brasileira, afastando-a, portanto, do português de Portugal, de pronúncia muito consonantal. (Cf. a pronúncia brasileira \*pi.neu, \*a.di.vo.ga.do, \*ri.ti.mo em lugar de pneu, ad.vo.ga.do, rít.mo). (Pessoa de Castro, 2015, p. 8)

Pessoa de Castro (2015, p. 9-10), ainda, em *Influência das línguas africanas no português brasileiro*, traz algumas das interferências das línguas bantu e iorubá na gramática do português brasileiro, em especial, o popular com destaque a seguir para os campos lexical, morfológico-sintático e o fonético-fonológico.

Segundo a autora, na língua portuguesa, o *campo lexical* é onde se encontram os dados mais robustos que provam a influência do contato de culturas e línguas, dentre elas, em maior quantidade as africanas. Algumas delas são:

1. Conservação de forma e de significados originais de palavras africanas tanto simples (samba, xingar, muamba, tanga, sunga etc) e compostas: lenga-lenga, Axé Opo Afonjá.
2. Sentido especial para palavras no português, seja por tradução direta de uma palavra africana como em mãe-de-santo (ialorixá), terreiro (casa de candomblé) ou por substituição de uma palavra considerada tabu na língua africana a exemplo de O Velho” (Omulu), e flor do Velho (pipoca).
3. Formação de palavras compostas em que há a presença de um elemento africano com um ou mais elementos do português, como em *bunda-mole, espada-de-ogum, limo-da-costa*.
4. Formação de palavras derivadas em português com base africana, a exemplo de *molecote, molecagem, xodozento, cachimbada, descachimbada, forrozeiro, sambista, encafifado, capangada, caçulinha, dengoso, bagunceiro*.

No campo *morfológico-sintático*, a autora destaca os seguintes elementos:

1. Marcação do plural apenas nos artigos que antecedem os substantivos como em *as casa* (as casas), *os menino* (os meninos), visto que era comum, na língua banta, o uso de prefixos para essa função;
2. Não marcação de gênero como acontece no Português Padrão, a exemplo de *minha senhor* (meu senhor);

Já no campo *fonético-fonológico*, considerando a estrutura silábica, a autora supracitada elenca:

1. Omissão/apagamento das consoantes finais das palavras fala (falar), dizê (dizer);
2. Transformação de consoantes em vogais como em Brasiu (Brasil) já que não existe sílabas que terminam com consoantes;
3. Tendência em desfazer encontros consonantais na mesma sílaba como em \*saravá (salvar) e fulô (flor) e, assim, produzindo uma sílaba nova;

Houaiss (1985, p. 91) afirma que “o português brasileiro nasce com diversidade”, vive e convive com ela, tanto regional como social e essa diversidade corrobora a fala de Castro (2001, p. 55), quando afirma que a língua dos escravizados no Brasil, com uma linguagem própria e mesclada do idioma natal e do português, fez contribuir para as alterações que hoje ainda são notadas, por exemplo, no fonetismo, no ritmo, no léxico e na sintaxe da fala popular dos brasileiros.

Na subseção, que segue, será discutido como alguns linguistas conceituam e explicam a hipótese da deriva linguística como a maior responsável para justificar as variações e mudanças no PVB no período colonial e pós-colonial. Nessa subseção, também, são suscitados outros pontos de vista sobre essa questão e explicações, igualmente, significativas que acabaram ganhando mais força entre os pesquisadores, tendo em vista a aceitação de uma avaliação dos aspectos sócio-históricos e das situações de contato, outro ponto a ser tratado mais adiante.

#### 2.2.2.1 Deriva secular das línguas românicas

Com o intuito de refletir sobre as origens do português brasileiro, Naro e Scherre (2007), apresentam algumas evidências históricas para a origem de algumas estruturas portuguesas que foram aos poucos ganhando espaço e se ampliando também em terras brasileiras. Os autores acreditam em uma confluência de fatores, de motivações para a ocorrência de determinados fenômenos linguísticos no Brasil que até então já tinham sido verificados em Portugal. “Se as sementes trazidas de lá germinaram mais rápido e cresceram mais forte, é que as condições, aqui, mostram-se mais propícias devido a uma CONFLUÊNCIA DE MOTIVOS<sup>8</sup>” (Naro e Scherre, 2007, p. 48).

---

<sup>8</sup> Destaque dos autores.

Naro e Scherre baseiam-se na noção de deriva linguística de Sapir (1949/1921). Sapir, 1949/1921, p. 150 *apud* Naro e Scherre (2007, p. 26) afirmam que “Language moves down time in a current of its own making. It has a drift.”<sup>9</sup> Isso quer dizer que as mudanças linguísticas acontecem com uma certa inclinação, ou seja, de alguma forma são previstas, embora sejam imperceptíveis no presente. É como se essas mudanças fossem resultado de um processo natural pelo qual a língua passa ao longo do tempo que lhe é próprio, prefiguradas, como o próprio Sapir descreve.

Os autores partem para explicar as suas bases teóricas à deriva secular,

É perfeitamente possível, entretanto, que a língua portuguesa já possuísse o embrião do novo sistema mais analítico, antes mesmo de sair da Europa. Tal estado de coisas se torna bastante plausível dada a deriva secular das línguas românicas, e indoeuropéias de maneira geral, em direção à uniformização morfológica, com a sobrevivência apenas das formas 'irregulares' mais salientes. (Naro; Scherre, 2007, p. 32).

Desse modo, para os autores, as mudanças que ocorreriam com o Português Popular do Brasil (doravante PPB) já estariam previstas no sistema visto que a gramática da língua já presumiria tais variações e/ou mudanças pela frequência de seus usos. Uma exemplificação dada pelos próprios autores que corrobora o processo da deriva é o fenômeno da queda do -s final no Brasil pois, diante de alguns fatos, foi possível levantar a hipótese de que esse fenômeno já tinha seu início no português dialetal europeu e estava apenas dando continuidade a uma deriva pré-românica. Naro e Scherre (2007) apontam que esse impulso do processo de perda de concordância nominal está ligado aos fenômenos fonológicos que foram trazidos da Europa e que se assemelham à concordância verbo/sujeito.

Os autores supracitados afirmam que é evidente que os primeiros africanos que foram trazidos ao Brasil falavam línguas africanas e, que em alguns casos e com distintos graus, começaram a adquirir conhecimento da língua geral e do português. Ainda asseguram que não há qualquer tipo de informação que seja capaz de provar que houve uma pidginização ou não das línguas africanas, no entanto, discorrem que:

Dado o fato de que as populações africanas eram muito concentradas geograficamente, com falantes de línguas banto para o sul e os da região do iorubá mais ao norte, parece inevitável a existência, entre os africanos no Brasil, de pelo menos duas línguas gerais de base africana, de forma semelhante aos sistemas de base tupi atestados. (Naro; Scherre, 2007, p. 30)

---

<sup>9</sup> “A língua se move ao longo do tempo num curso que lhe é próprio. Tem uma deriva”

Para os autores Naro e Scherre (2007, p. 31), é muito improvável que tenha existido uma língua *pidgin* ou crioula de base lexical portuguesa visto que já se tinha no território brasileiro outras “línguas gerais” não europeias, que atendiam às necessidades de comunicação da época.

Naro e Scherre (2007) fazem uma descrição geral de um modelo de desenvolvimento do PPB por eles proposto da seguinte forma:

- 1) A língua portuguesa falada em Portugal antes da colonização do Brasil já possuía uma deriva secular que a impulsionava ao longo de um vetor de desenvolvimento.
- 2) No Brasil, este vetor se encontrou com outras forças que reforçavam e expandiam a direção original.
- 3) No início, uma dessas forças era a pidginização, que exercia uma influência sobre o português através da língua geral tupi e da "língua de preto" europeia, revivificada no Brasil originalmente para uso com os ameríndios.
- 4) Ao longo de toda a história do Brasil, o processo de aprendizado do português como segunda língua teve seus efeitos documentados parcialmente.
- 5) Se existiu uma verdadeira língua crioula, caracterizada como sendo de "léxico português e gramática africana", ela cedo se evaporou sem deixar rastros na documentação. Sua possível influência no desenvolvimento do português do Brasil seria indistinguível da de outros eventuais pidgins ou crioulos de base não europeia. (Naro; Scherre, 2007, p. 46)

Diante do que foi apresentado sobre a deriva linguística, do ponto de vista dos linguistas Naro e Scherre, pode-se dizer que os autores rejeitam a hipótese de que tenham existido processos de reestruturação linguística significativos que estivessem ligados à crioulação prévia ou à interferência de substratos, por exemplo. Em suas hipóteses, também é rejeitado o fato de que qualquer outro mecanismo externo aos próprios contextos internos da estrutura linguística tenha contribuído para a formação do PVB. Vale destacar, no entanto, que os autores não ignoram

[...] a existência inquestionável de condições sócio-históricas normalmente propícias ao surgimento de línguas crioulas clássicas em terras hoje brasileiras: multilinguismo generalizado entre falantes adultos que, no início do contato, não partilhavam língua comum, em contexto de colonização e de consequentes relações linguísticas e sociais assimétricas. Também não duvidamos da provável existência durante o processo de aquisição do português por adultos de efeitos diretos de interferência de primeiras línguas em indivíduos ou até em comunidades relativamente isoladas de mesma origem. (Naro; Scherre, 2007, p. 118-119).

Desse modo, para os autores supracitados, as interferências a longo prazo não foram, visivelmente, evidenciadas diferentemente das estruturas que vieram de Portugal e, se houve algum tipo de efeito, foi passageiro ou não produziu resultado de *input* diferente do observado no europeu. Assim, consideram que os processos de variação/mudança existentes no PB não estão diretamente ligados com as línguas e povos que vieram para o Brasil, visto que já estavam presentes com a chegada dos colonizadores europeus, o que difere da hipótese de crioulização ou semicrioulização da língua portuguesa no Brasil.

Em um artigo de Mattos e Silva<sup>10</sup>, dedicado ao tema do PB para o Instituto Camões<sup>11</sup>, a autora posiciona-se em favor das teorias de contato, mas, ao mesmo tempo, afirma que a hipótese da deriva não deve ser descartada, pois elementos sincrônicos e diacrônicos se complementam para a tentativa de compreender a sócio-história do PB. Para a autora, o

Português Brasileiro heterogêneo apresenta variantes socioletais com configurações profundamente modificadas, que se aproximam dos crioulos de base portuguesa e variantes que se aproximam do Português Europeu. No primeiro caso, opino que uma *crioulização prévia, embora leve*, possa ter ocorrido no passado e, no segundo, a *deriva natural*, apressada pela história social do Brasil, não deve ser desconsiderada. Enfim, não se pode tratar como um conjunto homogêneo, unitário, o Português Brasileiro, nem numa perspectiva sincrônica, nem numa perspectiva diacrônica. (Mattos e Silva, s.p.)

Sobre essa perspectiva, Mello (2011) compartilha da posição tomada por Mattos e Silva, rejeitando os dois extremos - mudanças intrassistêmicas e deriva e a das mudanças intersistêmicas e inovações oriundas do contato linguístico -, e argumenta a favor da possibilidade de convergência entre as duas visões, a menos que sejam adotadas perspectivas múltiplas que levem a um equilíbrio entre as distintas hipóteses em relação a certos aspectos gramaticais do PVB. Para Mello (2011, p. 177), “um processo potencializaria o outro, ou seja, elementos gramaticais fracos, com tendência a apagamento, por exemplo, sofreria um incremento nesse processo, se constituíssem simplificações em um sistema adquirido como segunda língua”.

Mello (2011) afirma ainda que a sócio-história não pode ser desconsiderada no estudo da formação da uma língua. Para ela, o que deve ser feito é continuar investindo na compreensão de novos dados históricos através de estudos comparados, inclusive, entre falantes do PVB com outras variedades do português, não somente falada por indígenas, mas também por africanos

---

<sup>10</sup> Não há datação da publicação do artigo de Mattos e Silva no site do Instituto Camões.

<sup>11</sup> <http://cvc.instituto-camoes.pt/hlp/hlpbrasil/index.html>

nas ex-colônias de Portugal no território africano. Para isso, segundo a autora, os dados para novas análises devem ser consistentes a ponto de contribuírem para a compreensão e explicação do desenvolvimento do Português no Brasil.

Na subseção seguinte, serão mais bem discutidas as diferentes visões de alguns linguistas para o estudo das origens do PB, com foco na abordagem do contato linguístico, considerando a formação sócio-histórica no território brasileiro e, principalmente, nas questões que envolvem a transmissão linguística irregular ou semicrioulização.

#### 2.2.2.2 Contato linguístico

Nos estudos que envolvem a história do português brasileiro, por exemplo, muito se tem debatido se essa nova variedade do português se originou por um processo natural de evolução pelo qual a língua passou motivada por fatores internos cujas marcas encontradas no Brasil foram trazidas de Portugal, hipótese da deriva, ou se houve uma ruptura natural devido ao contato de outras línguas sobretudo das variedades mais populares, hipótese do contato. Pela hipótese da deriva, como já discutido na subseção anterior, Naro e Scherre (2007, p. 182) afirmam que “[...] no português do Brasil inexistente influência gramatical específica de qualquer língua africana” e, em oposição, a hipótese do contato, Lucchesi (2009, p. 28) argumenta que “não se pode pensar seriamente que a língua portuguesa não foi diretamente afetada pelo contato do português com as línguas africanas de uma forma bem ampla e representativa”. Nessa mesma linha, Mattos e Silva (2002, p. 456) destaca o elemento negro como o maior difusor do Português no Brasil e afirma que “essa população de origem africana – segmento demográfico dominante no Brasil colonial –, que teve de abdicar de suas línguas, pelas razões históricas conhecidas, e que adquiriu a língua de dominação, reformatando-a profundamente”.

Embora exista uma situação polarizada, dicotômica para a explicação sobre as origens do português brasileiro, Avelar e Galves (2014), posicionam-se a favor do contato linguístico e assumem o ponto de vista de Lucchesi e Baxter (2009, p. 101) quando afirmam que o que houve no Brasil foi um tipo específico de Transmissão Linguística Irregular<sup>12</sup> devido ao contato maciço entre falantes de diferentes línguas. Além do mais, não consideram a hipótese da crioulização na formação do PB. Os autores afirmam que

---

<sup>12</sup> Transmissão Linguística Irregular designa “[...] processos históricos de contato massivo entre línguas nos quais a língua do segmento que detém o poder político é tomada como modelo ou referência para os demais segmentos” e “[...] constitui um *continuum* de níveis diferenciados de socialização/nativização de uma L2 adquirida de forma mais ou menos imperfeita, em contextos sócio-históricos específicos” (Lucchesi, 2009, p. 109).

[...] certas marcas gramaticais singularizadoras do português brasileiro no conjunto das línguas românicas se devem à ação dos contatos interlinguísticos estabelecidos entre falantes de português e de línguas africanas (em particular, línguas bantas). O que estamos chamando de *ação* indica, em linhas gerais, duas situações diferentes: (a) transferência de padrões frásicos e propriedades morfossintáticas das línguas africanas para o português adquirido como segunda língua pelos africanos e (b) propriedades desencadeadas por dificuldades em reproduzir, por parte dos africanos, as marcas da língua tomada como alvo (no caso, o português) quando da sua aquisição. (Avelar; Galves, 2014, p. 243-244).

Exemplos comuns de marcas gramaticais singularizadoras do PB descritas por Avelar e Galves (2014, p. 244) estão:

- (i) construções de tópico sujeito, especificamente os casos em que os termos com interpretação locativa ou possessiva ocupam a posição de sujeito e concordam com o verbo;
- (ii) emprego de preposição *em* junto a complementos de verbos de movimento, bem como a variação entre ausência e presença de preposição na complementação de tais verbos;
- (iii) predicados dativos com duplo objeto;
- (iv) sintaxe pronominal, em particular no que tange à variação morfológica na forma dos pronomes em função da sua marca casual/função sintática;
- (v) nomes sem determinantes (os chamados *nomes nus*) singulares em posição argumental

Avelar e Galves (2014) debatem, assim, como as línguas africanas desempenharam um papel significativo na emergência da gramática de um português brasileiro, trazendo argumentos através de uma análise comparativa do ponto de vista morfossintático entre o português brasileiro, o africano e as variedades bantas. Desse modo, para os autores, a aquisição do Português como segunda Língua (L2) acarretou mudanças em duas direções. A primeira diz respeito às transferências de propriedades sintáticas e, a segunda, a reestruturação desencadeada devido à dificuldade no aprendizado de marcas gramaticais que eram específicas do português.

Ao apontar que o intenso contato entre os falantes do português e das línguas africanas, em especial, dos grupos bantos, gerou diferentes marcas na variedade brasileira, os autores explicam os dois fatores mais significativos. O primeiro está ligado ao fato de que a maior parte dos africanos no Brasil falavam um português mais próximo daquela variedade falada na África e, o segundo, o destaque de dois países africanos, Angola e Moçambique, que abrigavam povos do grupo banto possibilitando análises comparativas entre o PB e o Português em países africanos.

Petter (2011, p. 269) também sugere a inclusão de comparações do português com outras variedades de línguas africanas, em especial, daquelas que não desenvolveram crioulos como foi o caso do Português Angolano e do Português Moçambicano e, explica que,

[...] os estudos sobre o PA e o PM desfrutam de uma situação privilegiada de observação em que as diferentes línguas em contato ainda estão presentes, ainda são faladas e estão interagindo com o português. (Petter, 2011, p. 269).

Desse modo, o estudo dessas variantes pode fornecer dados importantes que também contribuem para o entendimento de especificidades do PB, do contato com as línguas africanas (do grupo banto) e das mudanças que estão em curso.

Lucchesi e Baxter (2009, p. 101), ao considerar que no Brasil houve uma espécie de Transmissão Linguística Irregular conceitua o termo “[...] para designar amplamente os processos históricos de contato maciço entre povos falantes de línguas tipologicamente diferenciadas, entre os séculos XVI e XIX, em decorrência da ação do colonialismo europeu”. Analisando o contexto da Transmissão Linguística Irregular<sup>13</sup> (doravante TLI), a língua do grupo dominante se impôs, a língua de superestrato, nesse caso, o português, “[...] de modo que os falantes das outras línguas, em sua maioria adultos, são forçados a adquiri-la em condições bastante adversas de aprendizado” e, com o surgimento das novas gerações, suas línguas de origem vão sendo abandonadas, criando assim, um lugar fértil para novos modelos de aquisição da língua portuguesa no Brasil.

Para os autores,

[...] tal processo de nativização da língua dominante ocorre de maneira irregular no sentido de que os dados primários de que as crianças que nascem nessas disposições dispõem para desenvolver a sua língua materna provêm praticamente de versões de segunda língua desenvolvidas entre os falantes adultos das outras línguas, que apresentam lacunas e reanálises em relação aos seus mecanismos gramaticais. (Lucchesi; Baxter, 2009, p. 101)

Naro e Scherre (2007) rebatizaram positivamente o termo TLI por Nativização pois acreditavam que a expressão dava a impressão de se tratar de algo anormal, errático e imprevisível. Nativização é tomado pelos autores “[...] no sentido estrito de passagem de uma língua não nativa a língua nativa de uma comunidade de fala [...] exclui o pressuposto de que a nativização é um processo instantâneo, como resultados uniformes em toda a comunidade” (Naro; Scherre, 2007, p. 53). Assim, diferentemente de Lucchesi e Baxter (2009), os autores

afirmam que a nativização não alterou estruturas variantes, mas que apenas alterou a frequência dos dados já trazidos pelos portugueses. Para Naro e Scherre (2007, p. 158),

O processo de 'transmissão linguística irregular' no Brasil, então, teve como entrada, ou estágio inicial, um estado de língua portuguesa em que já existiam as estruturas variantes, vistas hoje em dia, erroneamente, como exclusividade brasileiras [...], nenhuma estrutura nova foi criada pela 'transmissão linguística irregular' – ela simplesmente alterou, para cima, a frequência de certas estruturas variantes já existentes nos dialetos de menor prestígio social e que entravam apenas tangencialmente nas formas mais prestigiadas da língua.

Lucchesi (2009, p. 73) afirma que quando há uma

[...] grande concentração de africanos com pouco acesso aos modelos da língua portuguesa no interior do país, essas variedades linguísticas podem ter sofrido uma erosão gramatical e passado por processos de reestruturação, em níveis superiores aos que afetaram outras variedades do português brasileiro. (Lucchesi, 2009, p. 73).

O autor ainda afirma que, existiu, no Brasil, um processo mais amplo de TLI, caracterizada como do tipo leve e não de uma crioulição mais estrita, termo análogo aquele definido por Holm (2004) de *reestruturação parcial* ou ainda, *semicrioulo* por Silva Neto (1951[1963]). De acordo com Lucchesi,

[...] a aquisição precária do português pelos escravos trazidos da África e pelos índios integrados na sociedade brasileira e a nativização desse modelo defectivo de português como língua segunda nas gerações seguintes de seus descendentes endógamos e mestiços desencadearam um processo de transmissão linguística irregular que teve importantes consequências para a formação da atual realidade linguística brasileira, nomeadamente para as suas variedades populares. (Lucchesi, 2009, p. 71)

Lucchesi ainda pontua o fato de o negro estar integrado na sociedade, a exemplo dos ladinos, negros africanos, atuando nos serviços urbanos e domésticos, que já apresentavam uma certa proficiência, pois tinham maior acesso aos modelos de língua-alvo e adquiriam a língua portuguesa mais próxima dos seus senhores. Sobre esse grupo, Castro (1990, p. 104) afirma que

Ladinos e crioulos gozavam de grande prestígio sociológico na sua situação de bilíngues, isto é, na condição de poder participar de duas comunidades linguisticamente diferenciadas: a da casa-grande e da senzala. Enquanto na casa-grande eram os preferidos para os trabalhos domésticos, privando da sua intimidade, nas senzalas lhes era confiado o poder de disciplina e do comando

que se estendia às plantações e aos engenhos, na qualidade de capitães-domato e guardas pessoais de seus proprietários.

Um outro fator crucial para a não existência de línguas crioulas no Brasil foi o fato da sociedade brasileira ser bastante miscigenada, impossibilitando uma sociedade segregada racialmente, mas não socialmente (conforme ainda se vê, atualmente), cujos guetos dariam origem a uma formação de língua crioula e diferenciada já que esta resultaria da [...] “reestruturação independente da língua de superestrato, com forte influência do substrato”. (Lucchesi, 2009, p. 69). Assim, segundo o autor, o fato de a população brasileira ter uma população bastante integrada, principalmente de escravizados e ex-escravizados, a miscigenação serviu como um sentido de mão-dupla pois, ao mesmo tempo que os extratos mais baixos tinham seus falares introduzidos nas camadas médias e altas da população, eles também assimilavam os padrões linguísticos dominantes.

Nessa perspectiva, o que houve foi uma simplificação e/ou eliminação de algumas estruturas gramaticais ou diferenciação na frequência de alguns usos de formas não-marcadas, sendo base para a formação do PPB. Assim, conforme Lucchesi e Baxter (2009, p. 123), os processos para a explicação sobre as variedades populares do PB compreendem duas etapas: a primeira, *a fase inicial*, há uma

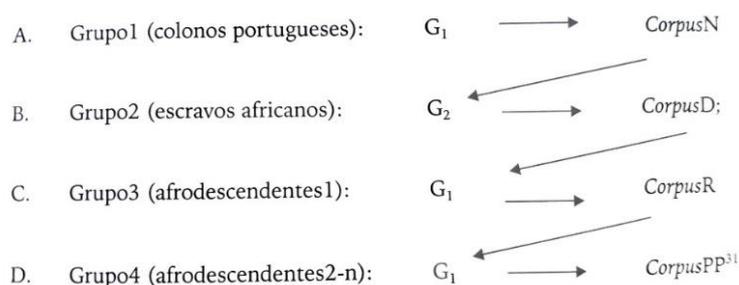
“variação, mais ou menos intensa, no uso dos mecanismos gramaticais, sendo mais rara a eliminação desses mecanismos; opacidade relativa nos estímulos-gatilhos que possibilitam a aquisição de valores diferenciados nos parâmetros sintáticos”. (Lucchesi e Baxter, 2009, p. 123).

Já, a segunda fase, denominada de *as fases seguintes*, mantém-se o quadro de variação, “em que se observa a concorrência entre as formas gramaticais reintroduzidas a partir da influência da LA<sup>14</sup> e potenciais processos de gramaticalização que emergem na heterogeneidade da fala; variação nas frequências de uso dos parâmetros sintáticos”.

A Figura 3, adaptada de Ian Roberts (2007, p. 390) pelos autores Lucchesi e Baxter (2009, p. 145) retrata um esquema de aquisição de uma língua em situação de contato, uma aquisição precária, imperfeita, a Transmissão Linguística Irregular. Nessa adaptação é possível entender o que aconteceu com a situação afro-brasileira.

---

<sup>14</sup> Língua-alvo, língua lexificadora ou língua de superestrato.

**Figura 3: Modelo de aquisição para situação de contato (TLI)**

Fonte: Roberts 2007, p. 390 apud Lucchesi e Baxter, 2009, p. 145 (adaptado pelos autores)

Tomando-se como base os elementos de A, B, C e D e suas relações, pode-se dizer que, na relação de A e B já possui uma aquisição defectiva de uma segunda língua por parte dos adultos africanos e o *input* é fornecido pelos colonos portugueses que são nativos. No caso da relação B e C, os filhos dos escravizados africanos já começam a falar a LP como nativos, no entanto, com um entendimento parecido, mas ainda precário que seus pais, anteriormente, receberam dos colonos portugueses. Nesse contexto, os filhos acabam reproduzindo o modelo recebido pelos pais de maneira também defectiva. Nas relações entre C e D, já é possível falar sobre uma transmissão linguística geracional pois, entre os grupos de afrodescendentes, a versão nativa do português já passou por reanálises caracterizando assim, um português reestruturado. Conforme os autores, os processos explicitados em A/B/C foram o tempo todo sendo replicados em solo brasileiro e, como houve várias levas de escravizados transplantados da África, sua duração não se prolongou por menos de 300 anos.

Para Lucchesi (2009), é bem possível que quando os falantes aprendem a língua-alvo partindo de um modelo defectivo e de forma precária, a reestruturação gramatical pode gerar uma variedade linguística muito diferente da língua de superstrato gerando uma **pidginização** e **crioulização**. No entanto, ele mesmo assegura que a TLI não gera, necessariamente, uma língua pidgin ou crioula, mas sim, uma formação de uma variedade histórica da língua-alvo, induzidos pelos contatos e que se caracteriza por apresentar processos de variação e mudança. Sobre o caso brasileiro, Lucchesi (2019, p. 241) apresenta dados sócio-históricos inibidores, alguns já apresentados anteriormente, que podem justificar no PB uma possível pidginização/crioulização nessa variedade:

- (i) Um percentual do grupo de falantes da língua de superstrato (30%) maior do que concebido como o máximo para que ocorra crioulização (20%).
- (ii) A grande representatividade de pequenos plantadores que possuíam de um a cinco escravos apenas, no conjunto de proprietários de escravos.
- (iii) Uma maior assimilação dos crioulos e sobretudo dos mulatos, que eram estimulados a adotar os padrões linguísticos e culturais da sociedade branca.

- (iv) O alto grau de mestiçagem presente na formação da sociedade brasileira.
- (v) O advento do ciclo do ouro, no século XVIII. (Lucchesi, 2019, p. 141)

Ribeiro (1995), em sua obra *O povo africano: a formação e o sentido do Brasil* (Ribeiro, 1995), mostra que na sociedade brasileira, com a participação indígena mais reduzida, o negro teve um papel decisivo na formação da sociedade, um agente da europeização que ajudou a difundir a língua de seu colonizador, que mesmo aprendendo a falar o português “aos berros do capataz”, acabou não somente, dominando a nova língua, como a refez, “emprestando singularidade ao português do Brasil, mas, também possibilitou sua difusão por todo o território” (Ribeiro, 1995, p. 220). Mattos e Silva (2004) acrescenta que

“É certamente no entrecruzar-se de variantes localizadas menos ou mais interferidas por marcas indígenas e/ou africanas, de variantes gerais mais ou menos africanizadas ou menos ou mais aportuguesadas que se definem e emergem traços característicos do português brasileiro, língua nacional.” (Mattos e Silva, 2004, p. 22)

Um outro ponto de vista para discussão sobre o contato linguístico no Brasil, Serafim da Silva Neto, ao entender que a variante padrão era ideologicamente prestigiada, afirmava o estudioso, em sua visão tradicional, Silva Neto (1960, p. 21),

“Por causa, precisamente, desta falta de prestígio é que a linguagem adulterada dos negros e índios não se impôs senão transitoriamente: todos que puderam adquirir uma cultura escolar, e que, por este motivo, possuíam o prestígio da literatura e da tradição reagiram contra ela”.

O autor acreditava na unidade da língua portuguesa no Brasil, não reconhecia os usos vernaculares, assim, para ele, a língua era a língua padrão. Toda a questão da precária e deficiente escolarização na sociedade brasileira também foi ignorada. Segundo o autor,

Se negros e índios falavam, nos primeiros tempos da colônia, um português deturpado, simplificado ao máximo em suas formas, era em virtude da condição social ínfima e de mínima cultura. Pouco a pouco, no contato com os brancos e ao sopro das luzes das escolas, negros, índios e mestiços de toda a espécie foram aperfeiçoando a sua linguagem (Silva Neto, 1950, p. 91).

Silva Neto (1963, p. 133 *apud* Petter, 2006, p. 132-133), admite que a influência africana, a linguagem simplificada e deturpada, se deu a depender da localidade, rural ou urbana, onde houve contração de mão de obra escrava e se havia acesso à escolarização. Assim, conforme Silva Neto, se a localidade estivesse situada em uma região de quilombo, ou se a escola não

fosse capaz de acabar com esse tipo de linguagem, estaria diante de uma língua crioula, “[...] uma simplificação extrema do português mal aprendido e imperfeitamente transmitido”.

Para Silvio Elia (1979), traços das línguas africanas não foram integradas na Língua Portuguesa pois a situação de contato não foi capaz de se fundir e nem gerar um crioulo no Brasil. Para ele, o que realmente aconteceu foi uma simples assimilação de uma cultura por outra, um semicrioulo, um “estágio preparatório” e não um crioulo propriamente dito. Elia (1994, p. 572), reafirma a posição de Silva Neto ao posicionar a língua portuguesa como a língua prestigiada

A [língua] de maior prestígio era, sem contraste, a portuguesa. As línguas indígenas e africanas, por ágrafas, não tinham a *Great Tradition*, a que se refere Joshua Fishman. O português quinhentista...fizera-se uma língua poderosa de uma nação poderosa. Nos contactos com a loquela inumerável de indígenas e africanos, a sua hegemonia seria fatal. (Elia, 1994, p. 572)

Gladstone Chaves de Melo (1946), na época, chama à atenção dos pesquisadores que a bibliografia sobre os estudos do português no Brasil, na seção III de seu livro “A influência africana” é escassa e que necessitava de mais estudos, mas, ele não deixa de defender a tese da unidade e conservadorismo do português brasileiro. Tanto Melo quanto Silva Neto colocaram a contribuição do negro na constituição da formação do Português no Brasil em segundo plano, como uma participação passiva, desacreditando da influência africana. No entanto, enquanto Melo desqualifica aspectos crioulizantes no território brasileiro, Silva Neto reconhece criouliização/semi-criouliização, que mais tarde, sofreu um processo de descriouliização no momento em que a ascensão dos mestiços aconteceu e que transformou o português padrão.

Guy (1981, 1989) defende que houve criouliização no passado e que os traços presentes no português vernáculo hoje, não foi devido à uma evolução natural e justifica a sua posição pelo estudo de dados morfossintáticos quando tratou da concordância de número, nominal e verbal e Holm (1987, 1992) segue a mesma linha de pensamento, no entanto, de semicriouliização pois, ao comparar os dados do PVB e línguas africanas e crioulas. Para Holm (1994, p. 59), é possível estabelecer um paralelismo de estruturas, “uma língua que resultou de um prolongado contacto entre uma língua crioula e não crioula”. Os autores tomaram por base a própria demografia brasileira que chegou a absorver uma média de 40% dos escravos que vieram transportados para o continente americano, com destaque para o período em que, no Brasil, a população afro-brasileira era superior a europeia.

Propostas como a de Guy e Holm têm suas importâncias para o estudo do PVB, e mesmo que alguns autores considerem que a existência isolada de algum tipo de criouliização no território brasileiro, elas têm sido enfraquecidas. Conforme já foi destacado aqui, fatos

impediram uma criouliização mais geral baseados em Lucchesi (2009, p. 70): a proporção entre a população de origem africana e branca, gerando um maior acesso à língua-alvo se comparado com situações de típicas de criouliização, o uso de línguas francas africanas como instrumento de interação dos escravos segregados e foragidos e o incentivo à proficiência em português.

Naro (1973) é um outro estudioso que criticou a criouliização, pois, segundo ele, preexistiu uma língua geral inibindo a formação de um crioulo além da ausência de uma mistura de línguas africanas de diferentes origens, já que grande parte dos africanos que se concentravam na Bahia era iorubá, enquanto, no restante do Brasil, a prevalência era banta. Fernando Tarallo (1993), também desconsidera a criouliização argumentando que, caso houvesse um crioulo de base portuguesa, o PB estaria se descrioulizando em direção da língua-alvo, a europeia, o que se contradiz com as pesquisas realizadas em que há o distanciamento com a língua de superestrato, o PE. Mello (2011, p. 178), afirma que “é muito provável que durante os primeiros séculos de colonização tenha havido uma grande instabilidade linguística em todo o então território brasileiro” visto que os primeiros contatos tiveram como base um léxico básico, muita gestualidade e, um posterior recurso na utilização de intérpretes a partir do momento em que as transações comerciais tornavam-se cada vez mais fortes no Brasil colonial. Essa fase mais inicial, pode se equivaler ao momento descrito por Mattos e Silva (2004) quando ela trata de um processo que se deu em um cenário de “multilinguismo generalizado” para um quadro atual de “multilinguismo localizado”, quando a política Linguística de Pombal é implantada e freia esse processo generalizado em meados do século XVII.

Pode-se concluir que os fatores sócio-históricos presentes no período colonial e no pós-colonial, baseados em uma situação linguística, em grande parte irregular, pelos falantes, bem como a ausência de uma normatização da língua falada e escrita, contribuíram fortemente para o surgimento de uma nova variedade que se distanciava do europeia, enquanto o PB vai se fortalecendo no território nacional.

### 2.3 VARIEDADES DO PORTUGUÊS EM ÁFRICA: PALOP

A Língua Portuguesa ocupa, atualmente, a quinta posição de idioma mais falado no mundo com cerca de 261 milhões de falantes<sup>15</sup> em 5 continentes, sendo língua oficial de nove países: Angola, Brasil, Cabo Verde, Guiné-Bissau, Guiné-Equatorial,

<sup>15</sup> [https://www.instituto-camoes.pt/images/eplp/Diptico\\_dlp16.pdf](https://www.instituto-camoes.pt/images/eplp/Diptico_dlp16.pdf)

Moçambique, Portugal, São Tomé e Príncipe, Timor Leste e Macau, pertencentes a quatro continentes, América, África, Europa e Ásia.

Dos nove países que adotaram o português como língua oficial, o Timor Leste é o único do território asiático e a Guiné Equatorial, o último estado-membro a integrar a Comunidade dos Países da Língua Portuguesa<sup>16</sup> (doravante CPLP) em 2014. Pode-se destacar ainda que em Angola, Cabo Verde, Guiné-Bissau, Moçambique e São Tomé e Príncipe e Guiné Equatorial, Países Africanos de Língua Oficial Portuguesa (doravante PALOP), são países de lusofonia africana que tiveram processos de independência de Portugal recentes, apenas em meados da década de 1970, após sucessivos processos de guerras e lutas contra exploração e domínio portugueses.

Os países que compõem a África Lusófona são países que têm se beneficiado de Protocolos de Cooperação Internacional para o Desenvolvimento, assim, firmam acordos nas diversas áreas, como a cultura, a educação a economia e a diplomacia. Na comunidade dos PALOP, propõe-se também a preservação da língua em seus territórios. Para dar sequência na luta comum que enfrentam seus membros, após conquistarem a liberdade, fóruns dos PALOP são realizados a fim continuar andando em direção ao progresso como também de propiciar o bem-estar que os seus povos precisam depois de quatro longos séculos de sofrimento e amarguras de um período colonial duro para os africanos. Buscam assim, dar longevidade aos laços fraternos e solidários entre seus membros.

Conforme Hagemeijer (2016, p. 49-50),

[...] como as variedades africanas do português ainda estão em uma fase de mutação mais intensa do que o português europeu e o brasileiro, reveste-se de especial importância a realização de novas recolhas que possam documentar a sua gradual estabilização. Um grande número de estudos sobre as variedades africanas do português procura assinalar tendências de divergência da norma do português europeu. Em muitos casos, a robustez destas tendências requer confirmação baseada em estudos quantitativos que também tenham em conta o perfil sociolinguístico dos informantes. (Hagemeijer, 2016, p. 49-50)

Na imagem a seguir, podem-se ver os países que compõem essa África lusófona.

---

<sup>16</sup> Com sede em Portugal, a CPLP foi criada em 1996 com o objetivo de promover o intercâmbio cultural, a defesa da Língua Portuguesa e a cooperação mútua entre seus membros através do diálogo cultural e regida por princípios.

**Figura 4: Mapa dos países que integram os PALOP**



Fonte: [https://pt.wikipedia.org/wiki/Pa%C3%ADses\\_Africanos\\_de\\_L%C3%ADngua\\_Oficial\\_Portuguesa](https://pt.wikipedia.org/wiki/Pa%C3%ADses_Africanos_de_L%C3%ADngua_Oficial_Portuguesa)

Apresentam-se, nesta seção, também, outras variedades de LP considerando o curto período de pós-independência do domínio português em países que integram os PALOP surgidas em decorrência das expansões ultramarinas no século XVI, com foco nas variedades africanas traçando um panorama sócio-histórico e linguístico. Antes de continuar essa discussão, serão feitas considerações sobre o português brasileiro, pertencente ao CPLP, sua formação em território americano e que se destaca em quantidade de falantes de LP e pela grande quantidade de estudos e pesquisa realizados sobre o tema.

Quanto aos países dos PALOP, **Angola**, é o segundo maior país de língua oficial portuguesa. Os primeiros portugueses, a exemplo de Diogo Cão, chegaram ainda no século XV e, algum tempo depois, o país se tornou uma colônia portuguesa. Luanda, sua capital, serviu de feitoria, ponto de apoio para embarques e abrigo de angolanos, chamados de “peças”, que seriam enviados pelos navios negreiros para o serviço escravo no Brasil. Alcançou sua independência de Portugal só em 11 de novembro de 1975. O país tem uma vastidão de recursos minerais e petróleo. Principalmente, após o término da Guerra Civil em 2002. Angola tem uma economia em ascensão, no entanto, o território ainda apresenta uma grande desigualdade econômica, já que a renda está concentrada nas mãos de uma minoria. A mortalidade infantil e a expectativa de vida não são boas. É um país que apresenta uma grande diversidade linguística e o português tem convivido com diversas línguas africanas, a exemplo do Kimbundo, Kikongo e do Umbundo.

Assim como no Brasil, Angola também vivenciou o contato massivo dos locais com o português vindo da Europa. Os locais em sua maioria pertencentes a tribos bantu, por mais que falassem livremente línguas aparentadas em seus territórios antes da chegada dos europeus, através do contato mútuo com eles, tiveram suas línguas impostas a outras regras.

Quando línguas ininteligíveis se encontram em um mesmo ambiente, há a possibilidade da língua-alvo sofrer reduções e simplificações, a princípio que é seguido de uma recomposição de elementos afetados, mesmo que em intensidades distintas. Nesse contexto de aprendizado de L2, as crianças que nasceram nessas situações acabaram desenvolvendo uma língua materna, como o próprio Lucchesi (2009) afirma, de versões de uma segunda língua que apresentam lacunas, que sofreram reanálises de adultos que são falantes de outras línguas, não sendo assim, uma transmissão geracional típica ou normal das línguas humanas.

Nas subseções mais adiante dessa seção, a realidade sociolinguística de Angola será mais bem apresentada e discutida, o que é um dos objetivos desse estudo.

**Cabo Verde** ou República do Cabo Verde é um país insular composta por 10 ilhas vulcânicas e uma área continental de 4.033 km<sup>2</sup>. Até o século XV, a ilha era deserta e após a chegada dos exploradores portugueses a ilha foi colonizada. As ilhas, por terem uma localização privilegiada virou uma economia atrativa para mercadores que tinham no comércio de escravos a sua fonte de renda nos séculos XVI e XVII. A ilha tem suas atividades escravocratas reduzidas no século XIX e só alcançou a sua independência em 1975, assim como Angola. Muito embora Cabo Verde tem a agricultura como sua principal riqueza, o país é um dos mais áridos e secos dos PALOP. No ano anterior à independência, houve um acordo do PAIGC<sup>17</sup> com Portugal, o que permitiu o direito ao povo à autodeterminação e independência, um período de transição, para uma futura conquista da independência do país. O português é a língua oficial e administrativa, das mídias e da alfabetização, no entanto, cotidianamente é bastante comum ver as pessoas falando o crioulo caboverdiano, ao norte na Ilhas Barlavento e ao sul, nas ilhas de Sotavento. Sobre essas questões da institucionalização da LP no território de Cabo Verde, o linguista Manuel Veiga (1994, p. 257) afirma que:

[...] se todo o nosso povo fala e compreende o Cabo-verdiano, a maioria, contudo, não fala o português. Apenas os nossos letrados, os que estudaram ou estudam são verdadeiramente bilingues. Na generalidade, o nosso povo tem um nível razoável de compreensão do Português, mas compreender uma língua não significa falar essa mesma língua. (Veiga, 1994, p. 257)

A imposição da língua portuguesa, deste modo, reforça a ideia de que a língua do superestrato sempre serviu de instrumento de dominação e de interesses políticos e econômicos.

---

<sup>17</sup> Partido Africano da Independência da Guiné e Cabo Verde.

Seu uso é assim justificado por uma pequena parcela detentora de poder, para que o país tenha visibilidade não só em caráter nacional, mas também internacional.

A língua nacional para a população em geral, até para quem mora fora do país, é o crioulo caboverdiano. Por outro lado, vale pontuar também a falsa ideia de que se um determinado povo vive naquele lugar, ele precisa se subjugar e aceitar os mandos e os desmandos da classe dominante, independentemente da sua vontade. No ideário daquele que é dominado, há a busca incessante pela preservação dos interesses coletivos, da língua que agrega, que une e que possibilita pertencimento naquela/daquela cultura e naquele/daquela lugar.

Em Cabo Verde, a LP é mais bem utilizada como uma ficção da homogeneidade de uma língua que destoa da população que é marcada pela diversidade e pluralidade linguística. Veiga (2009), em *o crioulo e o Português em Cabo Verde* afirma que:

Não obstante, o estatuto linguístico das duas línguas continua sendo muito diferente. Com efeito, enquanto a Lp é língua oficial e do ensino, da literatura, dos mídia e das situações formais de comunicação, o Ccv é língua de comunicação na família, língua das tradições orais, principal suporte musical, numa palavra, língua da oralidade e das situações informais de comunicação. Esta situação fez com que a Lp fosse sempre considerada como língua de prestígio e o Ccv como língua de amizade e do coração<sup>18</sup>.

Veiga (2009) defende o ensino do caboverdiano nas escolas. O autor ainda discorre sobre o fato de que, embora haja o amplo uso e preferência pelos negros, mestiços e brancos a usarem o crioulo local, com a ampliação do ensino formal e da criação dos seminários/liceus, acabou criando um entrave para o desenvolvimento da variedade nacional e fornecendo assim, condições mais estruturais para a afirmação e desenvolvimento da LP no país. Por isso, ele afirma que as línguas têm contextos contraditórios e a competição entre elas é desigual tanto nos âmbitos político, cultural quanto no linguístico.

A aquisição da língua portuguesa em Cabo Verde, não muito diferente do ensino em outros países africanos de língua oficial portuguesa, acontece de maneira que exclui o contexto multilíngue em que muitos falantes se encontram. As crianças normalmente têm o primeiro contato com a língua caboverdiana e quando iniciam os estudos no ensino fundamental, têm a obrigação de aprender uma outra língua que não representa a cultura deles e que está inserida em um outro contexto sociocultural. Sobre essas questões, Maritza Peña Rosabal, então ministra da Educação e da Família e Inclusão Social de Cabo Verde, entre os anos de 2016 e 2020, em

---

<sup>18</sup> <https://sibila.com.br/mapa-da-lingua/o-crioulo-e-o-portugues-em-cabo-verde/2753> (Veiga, 2009) SIBILA: revista de Poesia e Cultura. ANO 24, abr. 2009)

entrevista concedida à Revista Voa Português<sup>19</sup>, no seu primeiro ano de mandato, falou sobre a importância de ensinar o português como língua não materna desde o período pré-escolar. O projeto ao qual defende seria introduzido aos poucos à medida que as crianças fossem avançando para as séries seguintes. A ministra cabo-verdiana afirmou que " a língua portuguesa é abordada como língua primeira de Cabo Verde, quando não é. Temos uma eficácia do sistema muito baixa, em que apenas 44 por cento das crianças que começam o primeiro ano finalizam o 12º em tempo" e, ainda admitiu ainda que a "duplicidade linguística afeta o processo, pois a língua materna é o crioulo, mas como língua instrumental de trabalho e de comunicação temos que fortalecer a língua portuguesa".

**Guiné-Bissau** tem 36.120 km<sup>2</sup> e mais de um milhão e meio de habitantes. Até o século XVII, integrava o reino de Gabu. O explorador europeu chegou à costa do país em 1446, na figura de Nuno Tristão e tinha interesse no ouro do país. Com o tempo, a exploração das terras foi se interiorizando a partir do século XIX. O país também serviu de entreposto de tráfico de escravos e sua independência só foi declarada em 24 de setembro de 1973 e, só reconhecida, oficialmente, quase um ano depois. Guiné Bissau tem uma grande diversidade étnica, como os fulas e os povos mandingas (norte e nordeste), os balantas (costa sul), e os manjacos (centro e norte), além de uma maioria mestiça e uma minoria cabo-verdiana. Em relação ao português como língua oficial, apenas uma minoria fala, enquanto a maioria da população é falante de um crioulo de base portuguesa, sendo assim, não é um idioma falado como L1, muito embora, nas escolas, o ensino da LP é classificado como de uma língua materna.

Em Guiné-Bissau muitas pessoas são oriundas de várias regiões, por isso, não diferente de outros países africanos. No país há uma variedade de línguas nativas, cerca de 20 línguas africanas. Guiné Bissau também retratou na sua história o silenciamento das línguas nacionais em detrimento da LP, e, com o crioulo não foi diferente. Por mais de 500 anos, fez parte do imaginário guineense, o crioulo como um português que não era bem falado, por isso, aqueles que falavam a língua portuguesa eram considerados civilizados.

Ao se aproximarem do séc. XX, já nas lutas pela independência, o crioulo foi uma ferramenta, inclusive, que ajudou a desvencilhar sua história do processo escravocrata, e o português passou a ser não tão bem-visto devido ao seu passado colonial. Mesmo diante desse novo cenário, por questões políticas, houve a insistência de oficializar a LP no território pelo Amílcar Lopes Cabral, líder da independência.

---

<sup>19</sup> <https://www.voaportugues.com/a/portugues-segunda-lingua-cabo-verde/3626880.html>

Deste modo, mesmo o português sendo considerada a língua de prestígio, ele não impediu o crescente uso da língua crioula. Enquanto a LP é falada apenas por pessoas escolarizadas, o crioulo é utilizado por mais de 70% da população de Guiné Bissau, sendo a L1 dos mais jovens, especialmente, na área urbana e as línguas africanas, mais utilizadas nas zonas rurais. Sobre a complexidade linguística no ambiente escolar, por exemplo, Pereira (2011, p. 12) afirma que:

Não é raro encontrar, nas escolas guineenses, o professor a socorrer-se do crioulo guineense, como facilitador das aprendizagens, ou até como simples instrumento de comunicação professor/aluno e/ou professor/professor, dada a existência de indivíduos de várias etnias com línguas diferentes (Pereira, 2011, p.12).

A aquisição da língua portuguesa em Guiné-Bissau, considerada uma L2 ou até L3 para vários dos seus habitantes, muitas vezes limita-se ao contexto de sala de aula, já que, quando os alunos se encontram fora do espaço acadêmico utilizam a sua L1. As dificuldades para aprender a língua portuguesa são muitas pelos alunos guineenses, pois eles são obrigados a frequentarem as escolas primárias que impõem o aprendizado da leitura e da escrita tomando a língua portuguesa como se fosse sua língua materna. Além da LP estar restrita aos contextos de sala de aula, a forma como ela é ensinada, acaba gerando também dificuldades para a assimilação de conteúdos de outras disciplinas escolares. Desse modo, o ensino da língua oficial é baseado na memorização e na repetição, caracterizando um ensino não libertador.

**Moçambique** (República de Moçambique) é um país localizado no sudeste do continente africano cuja capital, Maputo, é também considerada a maior cidade do país. Sua população é superior a 30 milhões de habitantes em uma extensão territorial de 801.590 km<sup>2</sup>. É o segundo maior país integrante dos PALOP. Nos primeiros séculos d.C., os povos bantu migraram para a região e, a partir de junho de 1975, depois de 4 séculos de dominação portuguesa, a sua independência aconteceu. Assim como Angola, o país sofreu com uma intensa guerra civil que só findou em 1992. Sua economia é baseada na agricultura, com um aumento crescente da área das indústrias. Em Moçambique, fala-se, oficialmente, o português e, ele é falado por mais de 50% da população como L2, mas, vale destacar que, também são falados em seu território o macua, o tsonga, o ndau, o chiabo e o sena.

Assim como em todos os países que integram os PALOP, a presença portuguesa, aportou em Moçambique na figura do colonizador Vasco da Gama em 1497 e, em 1530, houve a povoação na região central do país. Ressalta-se aqui que, anterior à chegada dos portugueses, já havia relações comerciais com os árabes, principalmente, na parte norte. Não diferentemente de Angola e do Brasil, com a política linguística do período colonial, a Igreja também utilizou

de seu poderio ideológico como forma de dominação deixando as línguas africanas (línguas bantu) à margem, de forma preconceituosa, assim, eram vistas como dialetos, enquanto a LP europeia era a variedade respeitada, valorizada, oficial e obrigatória.

A língua, nesse contexto, é evidenciada como um instrumento ideológico muito poderoso de dominação dos povos, daqueles considerados sub-humanos e que possuem uma fala “deteriorada”. Sobre a expansão colonial pelos europeus, Lucchesi (2015, p.18) afirma que “a sujeição e exploração de povos e civilizações se transformaram em missão divina de propagação da fé e conversão do gentio; e a imposição da língua do colonizador era indispensável nesse processo de dominação física e espiritual”.

Conforme Timbane, (2014, s.p)

O Português de Moçambique surgiu de contextos semelhantes com as de TLI, mas teve um destino diferente. Por essa razão não surgiram pidgins nem crioulos, quer dizer, os moçambicanos não precisaram de um código emergencial porque houve insistência no uso das LB. Não houve um contato direto entre os colonos e os escravos porque haviam capatazes (cipaios) que serviam de intermediários (intérpretes). Os capatazes eram moçambicanos e tinham a obrigação de aprender português por ser assimilados. Por outro lado os escravos só tinham que cumprir com as obrigações impostas sob tutela dos capatazes. As LB continuavam a ser utilizadas nas igrejas independentistas africanas.

O autor discorre, sobretudo, que a situação do Brasil no período colonial se diferencia de Moçambique com base em alguns fatores como a presença de estrangeiros diversos que não só os africanos, a maior influência da igreja católica assim como a dominação do português em território brasileiro foi mais forte do que as línguas nativas. Ainda afirma que, mesmo havendo condições propícias para que um pidgin fosse implantado no Brasil, o sistema colonial não permitiu que isso acontecesse. No entanto, há ainda a concordância pelo autor de que, embora foram encontradas características de crioulição em Moçambique, elas não foram concretizadas.

Sobre a crioulição, com base em Pereira (2006), Timbane (2014, p. 10-11) explica alguns motivos para a não ocorrência de crioulos em Angola e Moçambique, dentre eles, o fato de não existir o uso de línguas africanas em todos os contextos de uso na comunidade; a presença de uma grande quantidade de intérpretes bilíngues; a vasta extensão territorial somada a um número pequeno de falantes do português se comparado à quantidade de falantes de línguas nativas (bantas); as pessoas que viviam em áreas mais rurais e o pouco fluxo para outras áreas mais urbanizadas favoreceu o uso das diversas línguas autóctones pelos seus habitantes.

Para Timbane (2014), o processo de nativização<sup>20</sup> da LP em Moçambique, foi consequência de uma aprendizagem irregular por uma quantidade pequena de moçambicanos e “os poucos escolarizados começaram a simplificar algumas regras, influenciados de certa forma pelas LB criando assim a variedade moçambicana. Assim, à medida que o tempo passa e que a LP vai se consolidando, uma variante de prestígio não apenas socialmente como também político, o número de falantes do português moçambicano vai crescendo.

Conforme Couto

A **apropriação de língua** também é um tipo de contato. No caso da **aquisição de L1** pela criança, trata-se de um **contato indivíduo-comunidade**, esta última representada pela mãe, pelo pai e demais pessoas que convivem com a criança [...]. Se se trata de **aprendizagem de L2** pelo adulto, temos um caso de contato indivíduo-comunidade em que o primeiro já traz o conhecimento de uma L1. (Couto, 2021, p. 180; grifo do autor)

Essa configuração da LP em Moçambique hoje, é reflexo dos contatos de tempos pretéritos entre colonizadores e colonizados. Segundo Timbane (2014), uma colonização marcada por uma grande resistência por parte dos nativos pertencentes a variadas etnias, assim como encontrado em Angola, um país multilíngue, por toda extensão territorial de Moçambique acrescido do insucesso dos colonizadores no ensino da LP, pois muitos deles não tinham interesse em ensinar a nova língua bem como não eram profissionais capacitados para tal função.

Semelhantemente, ao que aconteceu em Angola, em Moçambique, durante o período colonial, apenas uma parcela muito pequena da população podia frequentar as escolas, os filhos dos assimilados. O africano que adquiria costumes mais europeizados e que tivesse abandonado os seus usos religiosos, linguísticos de origem e costumes mais tradicionais, como por exemplo, ter aprendido a falar, ler e escrever bem em língua portuguesa, ter boas condutas, fazia dele um candidato com *status* “civilizado” a usufruir de certo prestígio na sociedade colonial da época.

Como já mencionado, em Moçambique, o número de falantes do português é cada vez mais crescente, como segunda língua, atingindo aproximadamente 11% da população nos dias atuais. Hagemeyer (2016, p. 47) afirma que a forma como se deu a presença portuguesa nesse país, de alguma maneira, ajuda a explicar “[...] o crescimento mais lento do número de falantes do português L1/L2 em Moçambique” pois o contato do colonizador/colonizado “[...]foi menor no tempo colonial e o conflito armado que se seguiu à independência foi menos desestruturante”

---

<sup>20</sup> O termo nativização adotado pelo autor tem por base as definições em Firmino (2002), Vilela (1995) e Lopes-Miguel (2004), sendo nativização “processo de transformação da norma-padrão europeia em PM, uma variedade que na base das LB adapta, integra na língua seus valores culturais, sua identidade, seus símbolos, seus objetos materiais de tal forma que seja sentida como pertence dos moçambicanos.

do que em Angola, por exemplo. Hoje, as línguas pertencentes à família de línguas bantas coexistem com o português.

Gonçalves (2001, p. 979) afirma que o PE padrão se diferencia do PM atual muito mais pelos traços fonético-fonológicos, morfológicos e sintáticos. Como as alterações gramaticais são mais estáveis e sistemáticas, as variações lexicais possuem um caráter mais disperso se comparado aos níveis mais gramaticais de sentido mais estrito. Ainda de acordo com Gonçalves (2001, p. 978), o interesse pela investigação do PM só cresceu mesmo em meados de 1980, e muito pouco foi estudado na época colonial e logo que a independência do país ocorreu. Sobre isso, a autora ainda afirma que a difusão do português em Moçambique teve seu processo de valorização amplo e consequências diversas assim como a sua consolidação de uma língua de “estatuto social a nível nacional”. Gonçalves (2001, p. 977) afirma que

O Português declarado << língua oficial>> e <<língua de unidade nacional>>, tornou-se verdadeiramente uma língua de prestígio: as novas gerações das classes mais favorecidas dos centros urbanos já não aprendem línguas bantu, podendo considerar-se que existe uma comunidade de locutores que não só se comunica exclusivamente em Português (mesmo que esta não seja a sua L1), como escolhem esta língua como a transmitir às novas gerações. (Gonçalves, 2001, p. 977)

Analisando o português não europeu, que mais tarde se caracterizaria como um português pluricêntrico, segundo Ivo Castro (2004, p. 84), foi marcado pela “história da ocupação do território, a formação do estado e os grandes movimentos da nação”. Pode-se tomar como observação as características de outras variedades do português no Brasil e o português falado em países da África.

O contexto linguístico em Moçambique, por exemplo, assim como em outras regiões africanas com presença histórica portuguesa, apresenta uma batalha constante para a preservação de suas línguas originais, mesmo com presença do português europeu como língua oficial nas escolas. A luta pela manutenção das línguas locais evidencia questões culturais, identitárias e sociais, sendo que a colonização empurra os colonizados a uma zona de não-ser, nem se representando como povo com base na sua cultura original, nem sendo aceito pelos colonizadores como iguais. De acordo com Fanon (2008, p. 26-34),

Há uma zona de não-ser, uma região extraordinariamente estéril e árida, uma rampa essencialmente despojada, onde um autêntico ressurgimento pode acontecer. [...] todo povo colonizado, isto é, todo povo no seio do qual nasceu um complexo de inferioridade devido ao sepultamento de sua originalidade

cultural, toma posição diante da linguagem da nação civilizadora, isto é, da cultura metropolitana.

Os moçambicanos, assim como outros povos colonizados por Portugal, vivenciaram um processo de exílio cultural e perda de identidade como resultado desta zona de não-ser, que é um legado da brutalidade colonizadora. O objetivo do discurso colonial, segundo Bhabha (2003, p. 111), “é apresentar o colonizado como uma população de tipos degenerados com base na origem racial de modo a justificar a conquista e estabelecer sistemas de administração e instrução”.

Há, então, de se examinar os diferentes contextos sócio-históricos para a formação da língua portuguesa em cada um de seus territórios, com suas diferenças, embora os laços estruturais existam. Segundo Mattos e Silva (2008, p. 4),

quando se fala de língua portuguesa una, na diversidade de suas manifestações, tem-se em mente - por um lado - a estrutura comum que está na base das suas diversificadas realizações e por outros os factores históricos que a unem e a definem como tal. (Mattos e Silva, 2008)

A abolição da escravatura, o fim do tráfico negreiro associados, a independência, a industrialização e a conseqüente urbanização favoreceram alguns novos contornos nas sociedades brasileira e africana, de formas diferentes e tempos diferentes, fizeram gerar um êxodo rural com uma grande quantidade de pessoas que passaram a viver na miséria, ao redor dos principais centros urbanos. Os migrantes recém-chegados nas grandes cidades levaram consigo um modo de vida mais rural, a cultura e língua do campo, ao mesmo tempo que acabavam sofrendo com a influência cultural da zona urbana e dos meios de comunicação de massa. Assim, com o fenômeno da urbanização, grandes contingentes da antiga população rural passaram a ser consumidores e moradores em um espaço urbano que lhes ofereciam uma variedade de língua mais socialmente valorizada e, em contrapartida, o falar dessas comunidades, frutos do êxodo, eram fortemente marginalizadas. Conforme Bortoni-Ricardo (2005), onde antes existia uma ampla variação diatópica (rural/urbano), passou a ter uma grande e profunda variação diastrática,

Os vernáculos ficaram por muito tempo mais ou menos circunscritos às regiões interioranas e isoladas. No século XX, assistimos, porém, a dois fenômenos de notáveis conseqüências linguísticas: a migração das populações de pequenas cidades e zonas rurais para os grandes centros e a difusão dos meios de comunicação de massa. Instala-se, então, nesses centros, um processo de diglossia, onde atuam duas forças antagônicas: por

um lado, o padrão tradicional de redução flexional da própria língua, exacerbado pela situação de contato entre dialetos diferentes; por outro, a pressão do prestígio da norma culta, imposta pela ação da escola, dos meios de comunicação e do status das classes mais favorecidas. (Bortoni-Ricardo, 2005, p. 33).

É no cenário de pluralismo étnico e cultural que as variedades da língua portuguesa se constituíram como línguas cheias de diversidades dialetais resultantes do contato com outras línguas, em que coexistiam diferentes povos e culturas. Esse intenso contato linguístico contribuiu para que a língua portuguesa lusitana sofresse modificações no território brasileiro, mas, e, em países africanos, desse origem a um português, peculiar, com traços distintos daquele utilizado em Portugal, o PB, o PA, PM, entre outros.

Vale destacar também, que alguns estudos, como é o caso de Gonçalves (2010), apontam para a evidência de variação visto que alguns usos não convergem com a norma europeia do português. Nesse estudo, a autora, ao tratar sobre a gênese do português de Moçambique, parte do pressuposto de que o português em terras moçambicanas:

está a emergir do processo da sua aquisição como língua segunda por crianças com línguas maternas bantu, num contexto em que nem sempre estão disponíveis amostras robustas do português europeu padrão, tomado oficialmente como norma de referência.” (Gonçalves, 2010, p.7).

As variedades africanas do português possuem traços linguísticos convergentes e divergentes do português europeu, por isso é importante explorar essas diferenças sem perder de vista o papel dos diversos contatos considerando os fatores sócio-históricos. Há ainda um desequilíbrio nas produções acadêmicas com maior presença de estudos envolvendo a variedade moçambicana, seguida da angolana, enquanto as variedades de Cabo Verde, São Tomé e Príncipe e Guiné Bissau ainda são menos exploradas.

**São Tomé e Príncipe** (doravante STP) – país localizado no Golfo da Guiné/Linha do Equador. A ilha de São Tomé, maior ilha do arquipélago, está a 140 km de Príncipe, e cerca de 250 e 225 km da costa noroeste do Gabão. Devido à sua localização, possui clima equatorial, quente e úmido e tem uma área total de 1001 km abrangendo mais de 200 mil habitantes. É o menor país de Língua Portuguesa e as pessoas da ilha falam o Português (língua oficial), o forro, o principense, o crioulo cabo-verdiano e angolar; inglês e francês como línguas estrangeiras. Foram ilhas desabitadas até 1470 e com a chegada de João de Santarém e Pedro Escobar, o território começou a ser povoado. No século XVI, as ilhas já tinham um comércio de escravos

oriundos de Angola, Moçambique e Cabo Verde, deste modo, a sua economia era dependente de escravos importados da África. No século XIX, a ilha introduz a produção do cacau e café em STP e vive os últimos dias de escravidão e os antigos escravos abandonam a ilha em massa. Até a sua independência (1975), vigora um sistema de roças, estilo casa grande/senzala que contribuíram para a construção de uma nova variedade da língua portuguesa.

Desde a fase de colonização, houve a tentativa de europeizar os povos locais através do ensino religioso pelos missionários portugueses. Mesmo assim, o forro continuou sendo a língua dominante até a sua independência. Não diferentemente de outros locais africanos, devido ao poderio português, as línguas crioulas passaram a ser consideradas pejorativas, e aqueles que desejassem ter algum tipo de ascensão social, proibiam os filhos também de utilizarem as variedades locais, inclusive, no seio familiar.

O português de STP se destaca, hoje, dentre os países da África que tem a língua portuguesa como língua oficial, pois ela é a única que tem o maior número de falantes tanto como L1 quanto L2 do português com mais de 98%, sobressaindo a L1 como língua oficial predominantemente falada pela população. Por outro lado, 36% declararam falantes do Forro, 8% do caboverdiano, 6% do Angola e apenas 1% do Lung'iê, conforme o Censo de 2012<sup>21</sup>. Deste modo, em nenhum outro lugar no continente africano, há um processo tão crescente da nativização do português, sendo considerado assim, um fenômeno particular. Conforme Gonçalves (2016, p. 26), há “um aumento gradual da erosão linguística dos crioulos, um processo tipicamente associado a contextos de línguas em contato, caracterizado pela perda de competência linguística da L1, motivada pela primazia do uso da L2”. Assim, no pós independência, atrelado ao processo de mobilidade social e da escolaridade, a língua do colonizador se tornou oficial e de maior prestígio, cada vez mais, se expandiu para uma grande parte da população como L1. Gonçalves e Hagemeijer, 2015, p. 92 concluem que

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

Tanto as variedades crioulas quanto a língua portuguesa em STP, contribui para o multilinguismo pois existe a convivência de uma L1 com o falar são-tomense tendo um certo grau de interferência; a língua oficial em período escolar e a língua que de fato é ensinada pela

---

<sup>21</sup> INE - <http://www.ine.st/demografia.html>.

grande parte dos professores já que muitos deles não tem a formação necessária sobre a língua oficial.

A **Guiné Equatorial** tem uma extensão territorial de 28.051km<sup>2</sup> com uma população de mais de 1.220.000 0 mil habitantes. O país faz fronteira com Camarões, Gabão e o Golfo da Guiné e, bem recentemente, oficializou a Língua Portuguesa em 20 de julho de 2010 como língua oficial e só 4 anos depois, tornou-se membro do CPLP. É também membro dos PALOP. Em 1494, os portugueses colonizaram as ilhas de Ano Bom e Corisco transformando-as em postos para o tráfico de escravizados. Vendeu muita mão de obra escravizada para a França e, mais tarde, para a Espanha e Inglaterra.

O Português se juntou a outras duas línguas oficiais, em primeiro lugar o espanhol, após o processo de descolonização e independência da Espanha e, em segundo, o francês. Em 1778, a Guiné Equatorial foi colonizada pelos espanhóis e a sua independência só aconteceu em 1968. Até hoje, na ilha de Ano Bom, ainda se usa o chamado Fá d'Ambô, uma língua crioula de base portuguesa que é bem parecida com o são-tomense de STP.

Uma das justificativas para a inserção da Guiné Equatorial no CPLP em relação à língua portuguesa refere-se ao fato de o país também ter vivido um passado colonial, em um primeiro momento no poderio dos portugueses e, séculos mais tarde, no poderio espanhol. Esse fato faz com que haja ainda muitas discussões e indagações sobre a oficialização de determinadas línguas, neste caso a LP, no pós-independência, já que na Guiné Equatorial, embora o francês e português sejam consideradas línguas oficiais, são ainda bastante desconhecidas e faladas como L1 pela população do país. A língua oficial de circulação é o espanhol, por isso, autores como Hobsbawn (2015) e Makoni (1998), Pennycook (2007), acreditam na invenção de tradição da língua visto que há uma ligação com o passado. Convém afirmar que a oficialização das línguas mais recentes no território seja vista como uma manobra para que possam também ter maior visibilidade em grandes blocos econômicos bem como para ser internacionalmente conhecido.

Quando algumas dessas “nações” europeias (Portugal, Espanha, Holanda, Inglaterra) começaram a dominar outras regiões do mundo, impuseram implícita ou explicitamente sua cultura e sua língua aos povos conquistados. Sempre que temos bilinguismo ou multilinguismo societário haverá uma língua dominante, em geral a língua do povo dominante. As demais línguas terão um status subordinado, frequentemente de **língua minoritária**, sobretudo se o número de seus falantes for menor do que o da língua dominante. (Couto, 2021, p. 114; grifo do autor)

Sobre algumas das produções acadêmicas que pesquisam sobre as variedades africanas do português, por exemplo, destacam-se os seguintes projetos: Os projetos *Panorama do Português Oral de Maputo* (doravante PPOM) e o *Corpus África* são exemplos de projetos, sediados no Centro de Linguística de Lisboa, que têm contribuído bastante para o estudo linguístico das variedades africanas do português.

O primeiro projeto, *PPOM*, está sob a coordenação de Perpétua Gonçalves e de Christopher Stroud desde o início da década de 1990. O *corpus* utilizado para análise é constituído de gravações orais de 100 falantes que moram em cinco bairros distintos na cidade de Maputo, capital moçambicana. Os colaboradores são de 5 faixas etárias e, para cada uma delas, 2 homens e 2 mulheres fazem parte. Há ainda a divisão dos falantes por nível de escolaridade E1, E2 e E3, que vão de características comunicativas mais limitadas até daqueles que possui nível superior completo ou incompleto.

O segundo, o *Corpus África*, é um *corpus* comparável de cinco variedades africanas do português e foi recentemente lançado pelo Centro de Linguística da Universidade de Lisboa (CLUL<sup>22</sup>). Nesse projeto, de *corpora* orais e escritos das variedades de Angola, Cabo Verde, Guiné-Bissau, Moçambique e São Tomé e Príncipe em relação à língua portuguesa. Embora seja um projeto mais novo, por agregarem *subcorpora* que podem ser comparadas entre si possibilita estudos e análises nos diferentes níveis linguísticos (lexical, morfossintático, fonológico e semântico-pragmático). O material de base de pesquisa é utilizado por diferentes áreas da Linguística, como a dialetologia e diacronia, filologia, gramática, fonética, fonologia e psicolinguística. Os *corpora* são compostos por materiais orais e escritos com homens e mulheres, alguns deles formais e, principalmente informais. Há ainda, o projeto das Variedades Africanas do Português (VAPOR), iniciado em 2008 com o intuito de coletar e trabalhar dados de fala tanto da L1 e L2 nos países africanos em que a LP é considerada língua oficial. Deste modo, é possível considerar que esse projeto tenha sido uma continuação prévia, em particular do *Corpus África*, ligados aos projetos de pesquisa da CLUL.

Dentre outros estudos, o PROFALA, um grupo de pesquisa coordenado pelas professoras Maria Elias Soares e Maria do Socorro Silva de Aragão na Universidade Federal do Ceará (UFC), conta com a ajuda da tecnologia da informação ao utilizar ferramentas e softwares com o objetivo de possibilitar diferentes análises sobre a variação linguística. Dentre um de seus estudos exploratórios, há o projeto de pesquisa “O Português falado nos PALOP e no Timor-Leste” em que o grupo investiga as variedades do português falado em Portugal e nos PALOP

---

<sup>22</sup> <https://www.clul.ulisboa.pt/>

- Angola, Cabo Verde, Guiné Bissau, Moçambique, São Tomé e Príncipe e no Timor-Leste, não se limitando apenas ao aprofundamento do português africano e suas diferenças com o PB, mas também discutindo questões que tratem sobre a atitude de seus falantes e suas representações sobre língua materna ou não, língua oficial ou nacional. Nesse sentido, as discussões também perpassam pela “reconstrução e redefinição do que vem sendo considerada “lusofonia”, a investigação em torno de políticas linguísticas de difusão do Português em Portugal e nos PALOP, à focalização das formas de Português faladas na África”<sup>23</sup>, sempre observando diferenças e semelhanças com o PB, em diferentes níveis da língua e utilizando as visões sociolinguística, geolinguística e discursiva para tais análises. O grupo de pesquisa supracitado desenvolve também outros subprojetos com alunos bolsistas da graduação (PIBIC/CNPq/UFC) como é o caso do projeto: O falar dos africanos e timorenses de língua oficial portuguesa – composição de corpus; com alunos do doutorado e pós-doutorado, a exemplos dos subprojetos: “Representações da Língua Portuguesa e construção de processos identitários no dizer de falantes de países africanos de língua oficial portuguesa (PALOPs) e do Timor-Leste” e “A situação linguística do Português nos países africanos de língua oficial portuguesa (PALOPs) e no Timor-Leste: as funções sociais e as políticas de difusão”.

O COMPARAPORT<sup>24</sup>, “Estudo comparativo de variedades africanas, brasileiras e europeias do Português” um desdobramento do “Projeto 21” faz parte da Associação de Linguística e Filologia da América Latina (doravante ALFAL). O COMPARAPORT é coordenado pelas professoras Sílvia Figueiredo Brandão (UFRJ) e Silvana Silva de Farias Araújo (UEFS). Entre os anos de 2011 e 2017, esse projeto de pesquisa, ocupou-se basicamente de estudos voltados sobre a concordância e, só após 2018, que o projeto amplia seus objetivos de estudo, passando a realizar estudos de diferentes perspectivas teóricas, de outras variáveis morfossintáticas, mas também do ponto de vista da fonética e fonologia, sempre em busca de parametrizações existentes nas variedades portuguesas. Segundo o projeto, através de uma análise contrastiva, é possível traçar padrões e diferenciação das variedades estudadas, observar tendências da mudança linguística, testar hipóteses for meio de fatores linguísticos e extralinguísticos, de hipóteses explicativas para os fenômenos observados bem como de contribuir para o ensino e elaboração de materiais.

---

<sup>23</sup><https://profala.ufc.br/pt/o-portugues-falado-nos-paises-africanos-de-lingua-oficial-portuguesa-palop-e-no-timor-leste/historico/>

<sup>24</sup> <https://www.mundoalfal.org/sites/default/files/proyectos/EstudoSS.htm>

O Corpora de variedades do Português em análise (doravante CORPORAPORT)<sup>25</sup> é um banco de dados que fornece materiais para muitas análises linguísticas, coordenado pelas professoras Sílvia Rodrigues Vieira e Sílvia Figueiredo Brandão da Universidade Federal do Rio de Janeiro (UFRJ). Enquanto a primeira desenvolveu os projetos “Padrões de concordância em variedades africanas, brasileiras e europeias do Português: produção e percepção” (CNPq – 2018-2020); “Contínuos de/em variedades do Português: análises contrastivas”; Gramática, variação e ensino: diagnose e propostas pedagógicas (2013-2018), a segunda, desenvolveu os projetos “Três Variedades Urbanas do Português em Contraste” (CNPq 2017-2019) e “O vocalismo átono no Português do Brasil e no Português de São Tomé” (FAPERJ 2015-2018). Em parceria, as duas pesquisadoras coordenam o grupo de trabalho Projeto 21- ALFAL, anteriormente citado.

O projeto Libolo<sup>26</sup>, sediado no município de Libolo, Kwanza-Sul (Angola) é um projeto que também desenvolve pesquisas ligadas à área da Linguística, da Antropologia e da História, sob a coordenação dos professores Carlos Filipe Guimarães Figueiredo da Universidade de Macau (China) e da professora Marcia Santos Duarte de Oliveira da Universidade de São Paulo (USP). Por ser um projeto internacional, seus pesquisadores são oriundos também de outras instituições brasileiras de ensino superior como a UFMG, UFPA, UnB, UFMA e UFRR com a parceria de outros institutos de ensino e órgãos administrativos do Libolo. Hoje esse projeto é reconhecido pela sua grande contribuição para os estudos da Linguística de Contato além de ser integrante da Cátedra da UNESCO que defende políticas de língua visando o multilinguismo.

Hagemeijer (2016, p. 50), em consonância com Petter (2011), afirma que

“a devida exploração destas diferenças (e semelhanças) entre as línguas com que o português está em contacto em África permite compreender melhor o papel do contacto face a processos mais gerais relacionados com a variação e mudança” (Hagemeijer, 2016, p. 50).

Assim, embora, todos os países falantes de língua portuguesa tenham sido afetados em diversos períodos pela colonização portuguesa, os resultados não foram os mesmos, tendo-se originado, inclusive, variedades mais modificadas, a exemplos das línguas crioulas; mas aqui será dada ênfase nas variedades da língua portuguesa. Tomando como exemplo o trabalho de

---

<sup>25</sup> <https://corporaport.lettras.ufrj.br/>

<sup>26</sup> <https://www.kalulo.com/projetolibolo>

Pissurno (2017) sobre o português de Maputo/Moçambique em relação os padrões de concordância, podem existir variedades dentro de uma variedade da língua. A autora afirma que existem posições intermediárias, ou seja, há um *continuum*, uma oscilação de comportamento entre os usos, e, esse fato confirma a hipótese de que a variedade moçambicana estaria ainda em processo de formação. Assim:

a posição do PM no meio do caminho, isto é, apesar de apresentar um nível bastante alto de preferência pela concordância padrão, seguindo seu modelo Europeu de (possível) aquisição, ele se afastaria do PE e se aproximaria mais das variedades parcialmente reestruturadas, como a brasileira, por exibir contextos mais variáveis de não aplicação da regra. (Pissurno, 2017, p. 199).

A autora além de afirmar sobre o processo de formação pelo qual o PM está passando, ela discorre ainda sobre a grande influência que as línguas locais exercem no português falado em Moçambique ao construir “sua identidade linguística, a partir dos meios plurais de aquisição de língua aos quais é exposta” (Pissurno, 2017, p. 199). Assim, é de se esperar, portanto que, tendo Moçambique uma variedade de português mais recente, ora se assemelhando ao PB, ora ao PE, o forte contato multilinguístico tem papel crucial nesse processo.

Desse modo, é fundamental considerar outros contextos e variedades, a exemplo das africanas, em que o português se fez língua oficial a fim de compreender melhor o cenário sociolinguístico desses povos na atualidade. Em termos linguísticos, já se pode verificar um distanciamento da norma europeia em relação às variedades não europeias, como as angolanas, moçambicanas e brasileiras. Segundo Petter (2007), é possível aventar a hipótese de haver um *continuum* afro-brasileiro do Português, pois, acredita-se que os contatos entre as línguas africanas e o português produziram consequências linguísticas semelhantes, por todas compartilharem o contato com línguas do grupo banto e, do ponto de vista sócio-histórico, resultarem fortemente de um processo de expansão da Língua Portuguesa decorrente do colonialismo português, alicerçada na mão de obra escravizada africana.

Feitas essas explicações sobre a formação e a transplantação da Língua Portuguesa, no território americano e africanos, é feita, na seção seguinte, uma exposição mais detalhada sobre a transplantação da Língua Portuguesa em Angola.

### 3 VARIEDADE ANGOLANA DA LÍNGUA PORTUGUESA: ASPECTOS SÓCIO-HISTÓRICOS E LINGUÍSTICOS

*Hoje  
somos as crianças nuas das sanzalas do mato  
os garotos sem escola a jogar a bola de trapos  
nos areais ao meio-dia  
somos nós mesmos  
os contratados a queimar vidas nos cafezais  
os homens negros ignorantes  
que devem respeitar o homem branco  
e temer o rico  
somos os teus filhos  
dos bairros de pretos  
além onde não chega a luz eléctrica.*

(fragmento de *Adeus a hora da largada*, Agostinho Neto.)

Apresenta-se nesta seção a variedade portuguesa em decorrência das expansões ultramarinas no século XVI, com foco na variedade angolana. Discutem-se sobre os troncos linguísticos e os povos autóctones africanos, suas subdivisões, suas classificações e uma breve descrição para cada um dos filios, com ênfase no tronco nigero-congolês, pois é onde estão situados os grupos de línguas bantas. Como os grupos bantos têm uma grande importância para a formação das línguas africanas e, conseqüentemente, na formação de variedades africanas do português, são elencados traços linguísticos comuns aos distintos grupos nos diferentes níveis da língua. Na sequência, buscam-se apresentar os grupos étnicos e as línguas bantas presentes em Angola e sua localização no território angolano. É traçado um panorama sócio-histórico e linguístico de Angola. Também é feita uma discussão sobre o fato de os angolanos minimizarem as marcas do colonialismo pós independência, ao passo que a territorialidade da língua portuguesa se concretiza e se torna um símbolo de unidade de uma nação emergente. Nesse contexto, um grupo de pessoas além de compartilharem os espaços, as culturas e as histórias, passam a apresentar um senso de pertencimento, de identidade nacional nas interações com os outros. As reflexões realizadas trazem algumas discussões sobre as políticas linguísticas para a implantação da língua portuguesa em Angola, as línguas autóctones e a situação de multilinguismo intenso em que o país se encontra.

### 3.1 CONHECENDO OS TRONCOS LINGUÍSTICOS EM ÁFRICA: AS LÍNGUAS E OS POVOS.

Como já mencionado neste texto, o empreendimento colonial levou a língua portuguesa para a África, desencadeando, nesse continente, desde os finais do século XV, diferentes situações de contato linguístico, ora possibilitando o surgimento de línguas crioulas – como ocorreu em São Tomé e Príncipe, em Cabo Verde e em Guiné-Bissau –, ora dando ensejo à reestruturação da língua portuguesa, como na formação das variedades angolana e moçambicana do português.

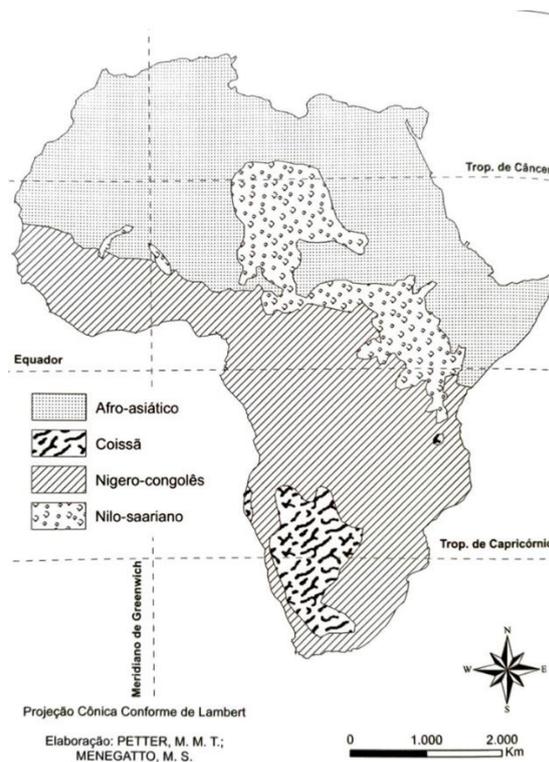
A “Era dos descobrimentos”, a partir do século XV, tornou-se um movimento impulsionador para a expansão da língua portuguesa que, até então, era mais restrita à área da Península Ibérica, sendo, de alguma forma, transplantada para outros continentes, como o americano e o africano. Esse movimento, segundo Castro (2004), foi responsável pela transfiguração radical da língua de Camões, pois, enquanto havia uma consolidação e reestruturação da língua portuguesa na Europa, “outras” assumiam contornos diferentes, a partir de uma história marcada de imposições culturais, políticas e linguísticas, dominações, sentimentos de superioridade, eurocentrismo, juízos de valor em relação ao outro e até silenciamentos de identidade dos povos envolvidos, em especial do negro africano.

É necessário entender a realidade linguística africana antes da chegada dos portugueses para uma caracterização da língua, em especial de Angola. Dos seis continentes do mundo, o africano é considerado um dos mais extensos do espaço geográfico abarcando 20% da área continental do planeta. É nesse espaço que abriga quase um terço das línguas faladas no mundo, uma média de mais de 2000 línguas africanas.

Considerando essa grande variedade de línguas africanas, Greenberg (1955), um linguista norte-americano, subdividiu as línguas africanas em 4 grandes troncos ou filios: o Nigero-congolês, o Afro-asiático, o Nilo-saariano e o Coissan, nas décadas de 1950 e 1960. Ele propôs a classificação com base no reconhecimento de certas semelhanças entre as línguas considerando hipóteses de relação genética entre as línguas africanas em seu estudo comparativo. A repartição criada por Greenberg é bastante importante para o estudo das línguas africanas visto que sua proposta foi a que melhor considerou todo o continente da África. Foi ele quem melhor descreveu uma metodologia para o estudo com os diferentes ramos das línguas, seus grupos e subgrupos, além de ser inspiração para os demais linguistas que o sucederam. Destaca-se, por exemplo, a classificação das línguas bantu por Ngunga (2004), africano, professor e pesquisador da Universidade Eduardo Mondlane-Moçambique (UEM), que tem realizado a descrição dessas línguas, em especial, de Moçambique. Sobre isso, o autor

não considera o Bantu uma família, nem indicativo de origem de línguas, deste modo, são do grupo bantu todas as línguas moçambicanas. Segundo Bonvini (2020, p. 23), a proposta dos troncos foi adotada por grande parte dos linguistas africanistas, com revisões em relação a organização interna dos troncos. No tocante à diferenciação de alguns aspectos internos dos troncos, Petter (2015, p. 51) descreve a proposta de Heine e Nurse (2000), mas também, em grande parte coincide com as famílias de língua já identificadas por Greenberg, conforme Figura 5:

**Figura 5: Línguas da África: troncos linguísticos, segundo Petter**



Fonte: Adaptado de Heine e Nurse (2000, p. 2 *apud* Petter, 2015, p. 52)

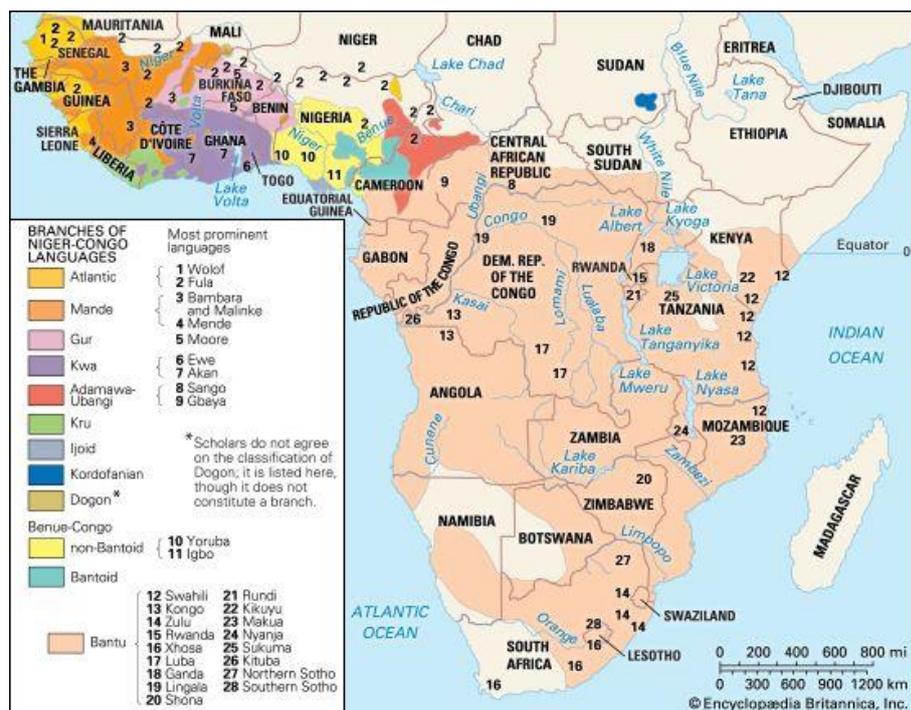
Baseado na proposta de Greenberg (1955) sobre o tronco *Nigero-congolês* (antigo nigero-cardofanês), e o que se vê na imagem 2, Bonvini (2020) afirma que esse tronco é o mais extenso em termos geográficos e abrange uma área que cobre da parte mais oeste do Senegal à parte mais leste do Quênia chegando à parte mais austral da África do Sul, com exceção do espaço coberto pelo tronco coissã.

O filo Nigero-congolês é constituído por 09 famílias linguísticas e, segundo Petter (2015 *apud* Lewis et al. 2014), conta com 1524 línguas além de possuir cerca de 470 milhões de falantes, dentre elas, falantes de línguas majoritárias da África, a exemplo de uólofe (Senegal),

iorubá e ibo (Nigéria), acan (Gana), entre outros. Ainda conforme a autora, “apesar de não ter sido feita ainda uma reconstituição do tronco como um todo, os especialistas dedicados ao seu estudo acreditam que se trata de uma verdadeira unidade genética e não apenas uma reunião de coincidências de traços tipológicos” (Petter, 2015, p. 54).

Considerando as famílias que pertencem ao tronco nigero-congolês, a benuê-congolesa é a que possui uma maior vastidão geográfica e a que tem a maior quantidade de línguas e falantes subdividindo-se em 11 subfamílias. O destaque é dado para a subfamília bantóide, subdividida em 2 ramos: bantóide do norte, que abrange a parte leste da Nigéria e oeste de Camarões), e a bantóide do sul, constituído de diversos subgrupos. Dentre os subgrupos, o banto está em evidência. Conforme se observa na Figura 6, há uma grande presença do grupo banto incluindo o sudoeste da Nigéria até as localidades ao sul do continente. Este grupo é o maior de todos os grupos linguísticos presentes na África, conforme imagem a seguir, e, também, o mais conhecido:

Figura 6: Mapa nigero-congolês (principais línguas) <sup>27</sup>



Fonte: <https://kids.britannica.com/students/assembly/view/18385>

<sup>27</sup> O marrom claro identifica as línguas da subfamília Bantu.

Guthrie (1967-1971) propôs uma classificação das línguas bantu que é bastante difundida e utilizada pelos linguistas e, mais tarde, foi atualizada por Maho (2003), segundo se vê na Figura 7, a seguir, há a presença das zonas e dos grupos de línguas bantu:

**Figura 7: Classificação das zonas geolinguísticas de Guthrie**



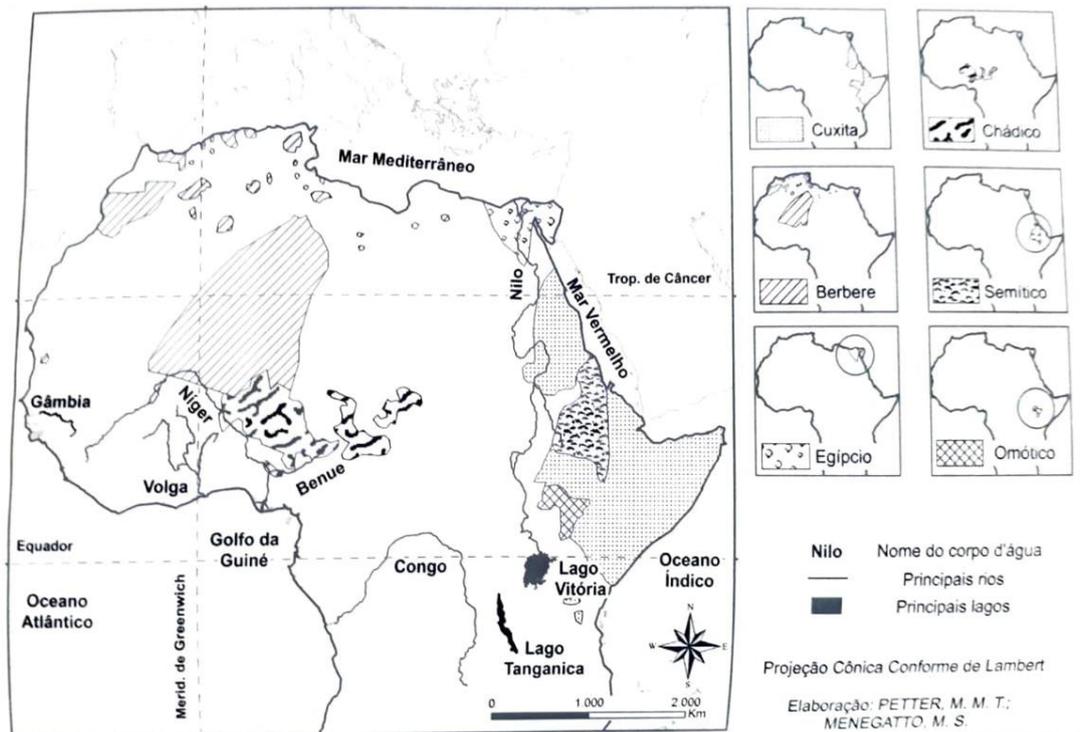
Fonte: <https://docplayer.com.br/86781906-Denominacoes-bantu-para-instrumentos-musicais-um-estudo-historico-comparativo.html>

A classificação de Guthrie se baseava em zonas geolinguísticas que foram identificadas por meio de letras e códigos numéricos de dois ou três dígitos com o objetivo de indicar a que grupo linguístico pertencia determinada língua. As primeiras, codificadas em letras maiúsculas (A, B, C, D, E, F, G, H, K, L, M, N, P, R, S), estavam as zonas onde as línguas eram faladas. Em sua classificação, cada zona foi dividida em vários grupos de língua com base na aproximação e distanciamento não só por motivos linguísticos, mas também geográficos. Os segundos, os números, em especial, os de dois dígitos, representavam o grupo linguístico a que pertencia e a língua propriamente dita. As línguas eram diferentes, mas tinham certos traços

linguísticos em comuns e que também eram consideradas próximas. Assim, os números indicavam uma quantidade de variedade de grupos nomeada por línguas representativas, e que ainda podiam ser acrescidos de uma letra minúscula que indicaria o dialeto de alguma língua específica.

Já o tronco *Afro-asiático*, anteriormente conhecido como hamíto-semítica (camito-semítico), passou a designar línguas bem antigas e que foram atestadas há mais de 5.000 anos, desde os primeiros escritos que se têm registro da humanidade. Nesse filo, conforme Lewis *et al.* (2014 apud Petter, 2015, p. 67), há a presença de 366 variedades linguísticas vivas, sem contar as línguas que foram extintas de que se tem conhecimento e pode ser mais bem visualizado na imagem a seguir, distribuídos nas seis famílias.

**Figura 8: Grupos linguísticos do tronco afro-asiático**



Fonte: Adaptado de Dimmendaal (2011, p. 209 apud PETTER, 2015, p. 66)

O filo afro-asiático compreende as línguas da África do Norte, e os 6 subgrupos são, com base em Petter (2015, p. 67-71), brevemente, caracterizados:

1. **Berbere** – representa 4 grupos de línguas e, como suas comunidades são mais dispersas e isoladas umas das outras, sua divisão é mais dialetal do que linguística (Maghreb);

2. **Chádico** – existem, aproximadamente, 140 línguas sendo que a mais conhecida e falada (como primeira e segunda língua) é a haucá, no noroeste da Nigéria.

3. **Egípcio** – língua que possui 4.500 de registros escritos e que permitiram reconstruir sua história até a sua morte, assim, diferentemente das demais, essa língua se submeteu à mudança linguística. O egípcio está presente na região do Nilo.

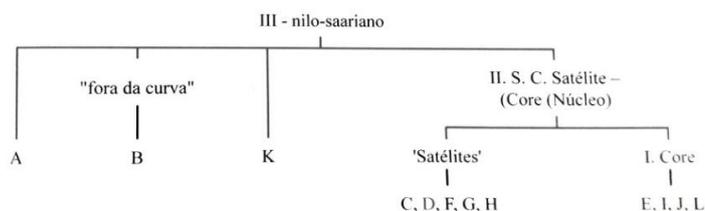
4. **Semítico** – é o mais estudado e mais bem compreendido. Tem um total de 50 variedades, entre línguas vivas e desaparecidas (já conhecidas). Uma parcela pode ser considerada “árabe” e possui, para a maioria dos estudiosos, 03 subfamílias (nordeste, noroeste e sul-semítico). Corresponde ao árabe e as línguas etíopes.

5. **Cuchita** – possui seis grupos de línguas com variedades que são muito distintas e, segundo critérios geográficos, estão classificadas em: *Cuchita Setentrional*, *Cuchita Central*, *Cuchita Oriental*, *Cuchita Oriental das terras baixas*, *Dulai* e *Línguas Cuchitas Meridionais da Tanzânia*, representadas pelo grupo iraco.

6. **Omótico** – composto por dois grupos principais de línguas: Omótico do Norte e o Omótico do Sul.

Um outro tronco descrito por Greenberg (1963) foi o *Nilo-saariano*, tronco considerado mais controverso pelos especialistas e o menos avançado no que se refere à classificação devido à um contato mais tardio com os europeus, bem como de poucos trabalhos que descreviam as línguas desse filo. Há, também, uma divergência quanto à quantidade de línguas visto que depende do que se considera língua e dialeto, mas, acredita-se que ele varia entre 195-198 línguas. Vale destacar, segundo Petter (2015), que a classificação proposta por Greenberg - de oeste para leste - foi, anos mais tarde, revista e modificada por Bender (1996-97) ao utilizar as letras de A-L para os níveis mais altos das famílias.

Diante de algumas tentativas que compreendessem melhor esse tronco linguístico na África, Bender (2000, p. 54 *apud* PETTER, 2015, p. 77), afirma que, embora seu estudo seja preliminar, há possibilidades de conseguir uma classificação genética mais apropriada, pois o autor acredita estar no caminho certo em sua investigação. Para isso, a metodologia empregada é “a da linguística histórico-comparativa, que se baseia no princípio de que a melhor evidência para agrupar as línguas é seu compartilhamento de *inovações*, isto é, itens que foram inovados num período comum de desenvolvimento, não itens *retidos* de um ancestral comum.” Desse modo, a classificação de Bender para o tronco Nilo-saariano é a seguinte:

**Figura 9: Classificação do tronco Nilo-saariano**<sup>28</sup>

**Fonte:** Bender (2000, p. 55 apud PETTER 2015, p. 77)<sup>29</sup>

Na literatura existem hipóteses de que o tronco Nilo-saariano faça parte de um tronco maior, em especial, o nigero-congolês, mas, vale destacar aqui que, embora muitos acreditem que o nigero-congolês e o Nilo-saariano possam ser um tronco, ainda não há unanimidade em associar o nigero-congolês ao ramo sudânico central. Esse fato reitera a necessidade de análises mais profundas que esclareçam os agrupamentos das línguas.

As línguas que fazem parte do tronco *Coissã* (Khoisan) constituem o menor dos 4 grupos de troncos que existem na África. Esse grupo de línguas é composto por apenas 24 línguas e se concentram mais na região da Namíbia e Botsuana. Apesar de ainda não ser possível determinar, precisamente, o número de falantes, estima-se que seja em torno de 200 mil. A maioria das línguas coissã já desapareceram, e devido às rápidas mudanças sociais, políticas e econômicas, acredita-se que grupos com menos falantes do coissã ainda estão ameaçados de extinção. Há o fato, ainda, de que essas minorias linguísticas sofrem com a estigmatização social. Sobre alguns dos traços linguísticos encontrados no grupo de línguas coissã, Petter (2015, p. 83) aponta que,

pode-se afirmar sem risco de engano que sistemas fonéticos coissãs estão entre os mais complexos do mundo e que são exclusivos da África; não só pela presença única de cliques, mas pela presença de raros traços tipológicos que se acrescentam a uma grande variedade fonológica. (Petter, 2015, p. 83)

Para um grupo de especialistas, o termo coissã não está relacionado a implicações genéticas, mas sim, utilizado para designar a todas as línguas que têm cliques não bantas e não cuchitas da África Oriental e Austral. As línguas coissã estão concentradas no deserto de Kalahari.

<sup>28</sup> A classificação de Petter (2015) foi adaptada de Bender (2000)

<sup>29</sup> A, B e K são considerados ramos independentes do Nilo-saariano e equivalem, respectivamente, ao sangai, saariano e culiaque. O S.C é o satélite-núcleo, o 4º ramo. Os ramos independentes do S-C são 6, a saber: C – maban, D – fur, F – sudânico central, G – berta, H - cunana e Núcleo. Este último, por sua vez, é formado por 4 famílias: E – sudânico oriental, I – coman, J – gumuz e L - cado.

### 3.1.1 As línguas do grupo bantu

Conforme já foi destacado na subseção anterior, o grupo *bantu* tem uma grande relevância na formação das línguas africanas e marcou uma das migrações mais intensas da história da África e, através do contato linguístico, pode-se compreender o porquê de determinadas línguas perderem ou ganharem alguns traços, mesmo não fazendo parte da mesma família ou do mesmo grupo.

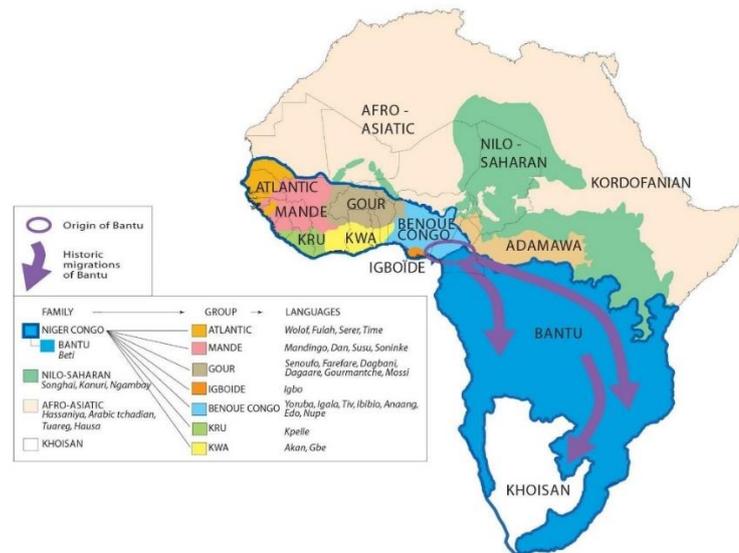
Segundo Sellier (2003, p. 9 *apud* Petter, 2015, p. 53), acredita-se que as línguas nigero-congolesas se iniciaram na região do alto Senegal e do Níger e, mais tarde, foi chegando às áreas mais ao leste e sudeste tornando-se um fenômeno maior chamado a Expansão dos *bantos*. De acordo com o autor, antes da chegada dos bantos, a região da floresta equatorial era povoada apenas pelos pigmeus, além disso, tanto as Áfricas oriental e austral eram dominadas por povos falantes das línguas coissãs. Tal cenário se modificou quando os bantos absorveram as línguas coissãs, restringindo-as as áreas mais sudoeste e ao extremo sul do continente africano.

O autor ainda pontua que o percurso da expansão dos bantos até a parte mais austral se deu pela margem norte da floresta equatorial e chegaram na região dos Grandes lagos ainda no primeiro milênio a.C. Para ele,

a expansão na floresta parece ter seguido os cursos d'água: os bantos teriam descido os afluentes da margem direita do rio Congo, depois subido este rio e seus afluentes da margem esquerda antes de atingir as savanas ao sul da floresta. A partir dessa região e a partir da África oriental, a expansão prosseguiu para o sul: os bantos chegaram ao sul do Zambeze no século IV e ao Limpopo, no século seguinte. (Sellier, 2003, p. 9 *apud* Petter, 2015, p. 53)

A figura 10, a seguir, apresenta o movimento migratório dos povos bantu para as regiões mais austrais do continente africano:

Figura 10: A Origem e as direções migratórias dos povos bantu



Fonte: [https://external-](https://external-preview.redd.it/SQNyxLCmY_4LMvbR3oKYVXiSjehCPqXkycBvfS3Jo_U.jpg?auto=webp&s=215c16427feb00862056a43fc2dd5a1ef13bec78)

[preview.redd.it/SQNyxLCmY\\_4LMvbR3oKYVXiSjehCPqXkycBvfS3Jo\\_U.jpg?auto=webp&s=215c16427feb00862056a43fc2dd5a1ef13bec78](https://external-preview.redd.it/SQNyxLCmY_4LMvbR3oKYVXiSjehCPqXkycBvfS3Jo_U.jpg?auto=webp&s=215c16427feb00862056a43fc2dd5a1ef13bec78)

O termo banto para as línguas africanas significava “pessoas”, “homens” cuja formação resultava do prefixo /ba-/, marca de plural e do radical /-tu/, /-ntu/ e variavam em outras línguas, como por exemplo, *banto* (quicongo), *wántù* (quissicongo); *àtù* (quimbundo). Bonvini (2020, p. 24), além de destacar o sentido para o termo banto, coloca em evidência o fato de que este termo foi utilizado pela primeira vez na Linguística, tecnicamente, por W.H.I. Bleek, em 1862, ao caracterizar uma família de línguas que utilizava o prefixo /ba-/ no plural, o que anteriormente já tinha sido definido pelo seu antecessor, H. Barth, a expressão *ba- languages* para fazer menção às culturas africanas que utilizam essas características. Na visão de Bleek (1862), o termo era considerado mais amplo, enquanto hoje, o termo é mais restrito e se aplica para as línguas da parte sul da África. Assim, até os anos 50, o termo banto designava uma família de referência (família plena e separada), como o nigero-congolês, e hoje, banto diz respeito aos membros do subgrupo bantóide do sul. Vale destacar ainda que, segundo Bonvini (2020), enquanto o termo banto para a Linguística é recente, historicamente, é algo que não está ligado a história da escravidão.

Considerando que as línguas bantas estão situadas no tronco do nigero-congolês, alguns traços linguísticos são compartilhados nos níveis fonológicos, morfológicos, sintáticos e, também do vocabulário comum. Para os seguintes níveis, algumas das características são elencadas por Petter (2015, p. 62 - 65):

Quadro 5: Traços linguísticos das línguas bantas

<p><b>Nível fonético-fonológico</b></p>	<p><b><u>Vogais:</u></b></p> <ol style="list-style-type: none"> <li>1. 05 vogais [i, e, a, o, u] ou 07 [i, e, ε, a, ɔ, o, u];</li> <li>2. Línguas podem ter 10 ou mais vogais – depende do avanço ou recuo da língua;</li> <li>3. Podem ser encontrados sistemas apenas orais e outros orais e nasais.</li> </ol> <p><b><u>Tons:</u></b></p> <ol style="list-style-type: none"> <li>1. Línguas são tonais (com algumas exceções, suailí, uólofe, entre outras);</li> </ol> <p><b><u>Estruturas silábicas:</u></b></p> <ol style="list-style-type: none"> <li>1. Tem restrições de sílabas fechadas (CVC);</li> <li>2. Poucas consoantes assumem a posição de coda;</li> <li>3. Na posição de ataque muitas línguas admitem apenas uma só consoante e ataques complexos do tipo NC [nasal, consoante], a exemplo de (<i>mb, nd</i> etc);</li> <li>4. Presença de nasais silábicas.</li> </ol> <p><b><u>Consoantes:</u></b></p> <ol style="list-style-type: none"> <li>1. Possuem 05 pontos de articulação – bilabial, dento/alveolar, palatal, velar e labiovelar;</li> <li>2. Labialização ocorre como articulação secundária;</li> <li>3. A palatalização é incomum e a velarização rara;</li> <li>4. Muitas oclusivas vozeadas pré-nasalizadas e, ocasionalmente, não vozeadas;</li> <li>5. Frequentes sistemas com fricativas /f, s/ e raros com /v, z/, /h/;</li> <li>6. Presença comum de / l, r, j, w/;</li> <li>7. Presença de línguas com consoantes foneticamente nasalizadas por formarem sílabas com vogais nasais.</li> </ol>
<p><b>Nível morfológico</b></p>	<ol style="list-style-type: none"> <li>1. Na morfologia nominal predominam os prefixos;</li> <li>2. Na morfologia verbal – distinções de tempo, aspecto e modo se realizam por meio de morfemas independentes, entre o sujeito e o objeto.</li> </ol>
<p><b>Nível sintático</b></p>	<ol style="list-style-type: none"> <li>1. tendência à ordem rígida da frase SVOX (sujeito, verbo, objeto e oblíquos - complementos adverbiais);</li> </ol>

	2. Poucas línguas são ordenadas em SOV (sujeito, objeto e verbo)
<b>Nível lexical</b>	<p>1. Possuem muitos cognatos para os itens reconstruídos</p> <p>2. lacunas de dados dificultam a reconstrução em larga escala, mesmo com a extensão e importância desse tronco;</p> <p>3. fragmentos das evidências morfo-lexical podem indicar uma unidade genética do tronco, de acordo com Williamson e Blench (2000, p. 41)</p>

Fonte: Quadro elaborado com base em Petter (2015, p. 62 - 65)

Na seção seguinte, são apresentadas algumas classificações realizadas sobre os diversos grupos linguísticos em solo africano, os grupos de línguas e as áreas de maior concentração dos povos na África, em especial, sobre os grupos étnicos presentes em Angola.

### 3.1.1.1 As línguas bantu e os grupos étnicos em Angola

Dentre os estados bantos em solo africano, um deles teve um grande destaque, o reino do Congo, que dominava grande parte do que hoje é conhecido como Congo e Angola. Considerando a classificação de Guthrie e a atualização de Maho, Inverno (2018), elenca as áreas, os respectivos grupos linguísticos e os povos angolanos de origem africana. Percebe-se, então, que em Angola, as zonas bantu que predominam são a H (10, 20), R (30,13, 20,10) e K (30, 10, 12), sendo que a capital, encontra-se na zona H com a língua representada pelo Kimbundo.

**Quadro 6: Principais grupos de línguas, povos e respectivas áreas dos povos angolanos**

<b>Origem</b>	<b>Povo</b>	<b>Língua</b>	<b>Províncias</b>
<b>Bantu</b>	Ambundu	Kimbundo ( <b>H20</b> )	Bengo, <b>Luanda</b> , Cuanza Norte, Cuanza Sul, Malanje
	Bakongo	Kikongo ( <b>H10</b> )	Zaire, Uíge (também no Bengo, Malanje, Cabinda, Cuanza Norte)
	Ovahellelo	Oshihellelo ( <b>R30</b> )	Namibe (também em Benguela, Cunene, Huíla)
	Ovanyaneka-Nkumbi	Olunyaneka ( <b>R13</b> )	Huíla (também no Namibe, Cunene e Benguela)
	Ovambo	Ovakwanyama ( <b>R20</b> )	Cunene (também no Cuando-Cubango)

	Ovandongua	Oshindonga ( <b>K30</b> )	Cuando-Cubango
	Ovimbundu	Umbundo ( <b>R10</b> )	Bié, Huambo, Benguela e Namibe (também no Cuanza Sul e Huíla)
	Tucokwe	Cokwe ( <b>K10</b> )	Lunda Norte, Lunda Sul, Moxico e Bié
	Vangangela	Ngangela ( <b>K12</b> )	Cuando-Cubango (também no Moxico)
Khoisan	Khoisan	Khoi ---	Huíla e Namibe
	Vátwa	Vátwa ---	

Fonte: Inverno (2018, p. 86). Destaque nosso.

Conforme se vê no quadro anterior, e fazendo uma relação com língua/povo bantu, tem-se o seguinte: a língua Kimbundo é falada pelos Ambundus; o Kikongo é falado pelos Bakongos; Oshihelelo é falado pelo grupo Ovahelero; os Ovanyaneka-Nkumbi falam o Olunyaneka; os Ovambos falam o Ovakwanyama; Ovandongua usam o Oshindonga, os Ovimbundu falam o Umbundo; os Tucokwe usam a língua Cokwe, os Vangangela utilizam o Ngangela. Há ainda, outras línguas que não pertencem à família banto como o Khoi falado pelo povo Khoisan e o Vátwa falado pela tribo que recebe o mesmo nome. É necessário reforçar que, similarmente, ao processo de colonização ocorrida no Brasil, as línguas dos povos autóctones têm sido “substituídas” pelo português, única língua oficial do país, e as línguas nacionais vêm perdendo a sua vivacidade, a sua vitalidade.

No mapa (Figura 11), a seguir, é possível ver de uma forma mais esquemática, a distribuição geográfica dos grupos étnicos presentes em Angola e as línguas faladas por eles:



conservem uma certa homogeneidade linguística como fruto da forte coesão cultural do grupo.

Conforme o Censo realizado em 2014, a língua angolana mais falada é o Umbundo (23%), seguida do quimbundo, do quicongo e do Cokwe. Esse mesmo Censo traz informações acerca das línguas faladas em Luanda, apontando o português como língua majoritária, no entanto, ele não mostra se esse domínio leva em consideração as condições em que o Português é falado, se como língua materna (L1), se adquirida (L2) ou de uma forma generalizada.

Esse fato impossibilita uma discussão mais aprofundada a respeito do quantitativo de falantes das línguas, se L1 ou L2. Luanda, até os dias atuais, ainda é marcada por intenso multilinguismo, onde diversas outras línguas estão convivendo com a língua oficial: a portuguesa. Como já relatado em outra subseção dessa tese, esse fato se deve ao movimento migratório provocado pela guerra de independência, pela procura por melhores condições de vida e trabalho, entre outros fatores sociais. Cabe salientar que Luanda desde a sua formação, já possuía uma diversidade muito grande de línguas e esse movimento fez com que esse número se intensificasse. A realidade linguística atual de Luanda é muito complexa. Na província em que a cidade de Luanda se encontra, há uma quantidade maior de falantes do Quimbundo.

Embora, como foi dito anteriormente, o português tenha o seu reconhecimento como língua oficial no país, algumas alternativas estão sendo colocadas em discussão e estudos estão sendo realizados na tentativa de conhecer e descrever melhor as línguas nacionais utilizadas em Angola. Ainda com base na entrevista realizada por Severo (2015, p. 9), alguns dos estudos realizados sobre as línguas nacionais são:

1. Estudo da morfologia e sintaxe da língua Ngangela, pelo professor dr. Zavoni Ntondo no Instituto de Línguas Nacionais com a proposta de criação de alfabetos atentando-se para as especificidades de cada uma das línguas nacionais já elencadas;

2. Pesquisa e ensino das línguas cokwe, umbundo, quimbundo e kikongo, línguas que têm sido também alvo de políticas de ensino nas regiões leste, sul, centro-oeste e norte do país pelo Departamento de Línguas e Literaturas em Línguas Angolanas da Universidade Agostinho Neto;

3. Iniciativas de difusão do ensino de línguas angolanas, com interatividade, nas plataformas digitais, como o projeto EVALINA<sup>31</sup>;

---

<sup>31</sup> <https://pt-br.facebook.com/ProjectoEvalina/>

4. Uso de línguas nacionais e do português angolano nas rádios do país, a exemplo da rádio Ngola Yetu<sup>32</sup>.

### 3.2 REPÚBLICA DE ANGOLA

A República de Angola está localizada na região austral, na costa ocidental do continente africano, possui um vasto território com uma superfície de 1.246.700 km<sup>2</sup>. Em termos comparativos, equivale a 5 vezes o tamanho do Estado de São Paulo. Angola é o segundo maior país de língua oficial portuguesa, o 7º maior país da África em extensão territorial e faz fronteira ao norte e nordeste com o Congo e a República Democrática do Congo, a leste pela Zâmbia, e a oeste está o Oceânico Atlântico e, ao sul, pela Namíbia, conforme se vê na imagem a seguir:

**Figura 12: Localização de Angola e países fronteiriços**



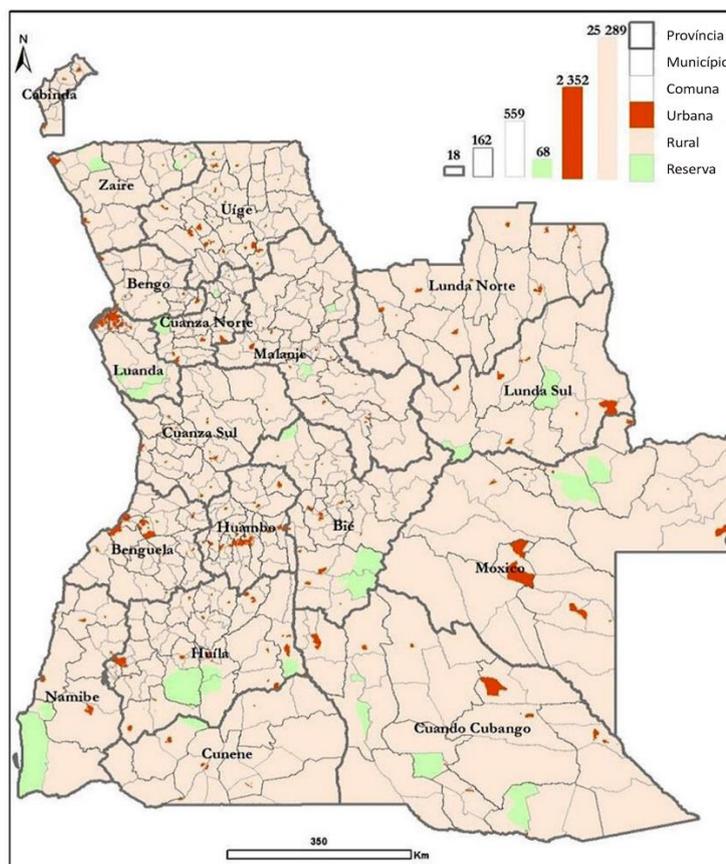
Fonte: [http://d2z7bzwflv7old.cloudfront.net/cdn\\_image/exW\\_1200/images/maps/en/ao/ao-area.gif](http://d2z7bzwflv7old.cloudfront.net/cdn_image/exW_1200/images/maps/en/ao/ao-area.gif)

A sua capital é Luanda, cidade mais populosa do país. De acordo com o Recenseamento Geral da População e Habitação (doravante RGPH), realizado em 2014, o primeiro censo pós independência, foi também o primeiro realizado utilizando padrões de qualidade internacional, diferentemente da realizada em 1970. O RGPH/2014 aconteceu entre os períodos de 16 e 31 de maio e teve uma cobertura de 94.4%. De acordo com o recenseamento, a população é composta por mais de 25 milhões de habitantes e estima-se que sua população cresça ainda mais. Em termos populacionais, não existe uma igualdade na distribuição dos seus habitantes no território, visto que a maioria da população se concentra nas áreas mais costeiras, ou seja, na

<sup>32</sup> <https://rna.ao/rna.ao/ngola-yetu/>.

parte mais oeste do território bem como nos maiores centros urbanos, enquanto 37.3% vivem em áreas rurais. Em termos administrativos, sua área é composta por 18 províncias que se subdividem em 162 municípios e em 559 comunas. No mapa que segue, são mostradas as suas distribuições no território angolano:

**Figura 13: Municípios, comunas e províncias angolanas (2014)**



Fonte: INE (2016, p. 27)

A seguir, no Quadro 7, há o destaque para as sete províncias mais populosas, a exemplo de Luanda, Huíla, Benguela, Huambo, Cuanza Sul, Uíge e Bié, equivalendo em mais de 70% de concentração populacional nessas áreas de acordo com os dados do INE.

**Quadro 7: Províncias de Angola em relação à população**

Bengo – 1.4%	Moxico – 2.9%
Cabinda – 2.8%	Lunda Sul – 2.1%
Cuanza Norte – 1.7%	Namibe – 1.9%
<b>Cuanza Sul – 7.3%</b>	Cuando Cubango - 2.1%
<b>Luanda – 27%</b>	<b>Bié -5.6%</b>
Lunda Norte – 3.4%	Cunene – 3.8%

Malanje – 3.8%	<b>Huíla – 9.7%</b>
<b>Uíge – 5.8%</b>	<b>Huambo – 7.9%</b>
Zaire – 2.3%	<b>Benguela – 8.6%</b>

Ainda conforme o documento, a população entre homens e mulheres é equilibrada, com um leve aumento para o sexo feminino equivalendo a 51.6%. Sobre a escolaridade<sup>33</sup>, percebe-se que a população ainda não é tão escolarizada visto que apenas 46.2% têm o nível primário, 19.8% o I ciclo do secundário e menos de 13% conseguiram completar o II ciclo do secundário. Os dados são mais alarmantes até mesmo quando se analisa a quantidade de pessoas que concluíram o ensino superior, bacharelado ou licenciatura, somando apenas 2.3% da população recenseada. Ainda há de considerar que 22.5% dos angolanos nunca frequentaram a escola e 32.3% já frequentaram alguma vez a escola, mas acabaram abandonando os estudos. Conforme o Censo realizado em 2014, a língua angolana mais falada como L2 é o umbundo com 23% e as línguas kikongo e quimbundo seguem com cerca de 8% cada uma, já, “o Português é falado por mais de metade da população (71%) com maior predominância nas áreas urbanas onde 85% da população fala a língua portuguesa, enquanto, somente 49% na área rural” (INE, 2016, p. 51). Mais detalhadamente, os dados podem ser analisados na figura 14, a seguir:

---

<sup>33</sup> O ensino em Angola é dividido entre os seguintes níveis de educação: pré-escolar (03 meses aos 05 anos de idade), o primário (da 1ª à 6ª classe), o secundário (da 7ª à 12ª classe) e o superior (até aos 6 anos de curso de licenciatura e aos 03 anos de bacharelado). Conforme o governo angolano, é obrigatório concluir a 9ª classe escolar.

**Figura 14: Línguas faladas pela população com dois ou mais anos pela área de residência (Zona rural, zona urbana e Angola)**

Línguas faladas	Número de falantes e correspondente percentual da população					
	Zona urbana		Zona rural		Angola	
Mais de uma	5.539.833	37,0%	3.480.571	39,7%	9.020.404	38,0%
Português	12.644.358	84,4%	4.246.383	48,5%	16.890.741	71,2%
Umbundu	2.502.897	16,7%	2.946.921	33,6%	5.449.818	23,0%
Kikongo	1.177.540	7,9%	778.651	8,9%	1.956.191	8,2%
Kimbundu	1.014.811	6,8%	841.140	9,6%	1.855.951	7,8%
Cokwe	1.006.165	6,7%	546.854	6,2%	1.553.019	6,5%
Nhaneca	182.515	1,2%	629.842	7,2%	812.357	3,4%
Ngangela	311.164	2,1%	427.906	4,9%	739.070	3,1%
Fiote	382.582	2,6%	185.714	2,1%	568.296	2,4%
Kwanhama	104.276	0,7%	433.257	4,9%	537.533	2,3%
Luvale	94.232	0,6%	153.769	1,8%	248.001	1,0%
Muhumbi	83.800	0,6%	419.081	4,8%	502.881	2,1%
Outras LN	494.778	3,3%	359.267	4,1%	854.045	3,6%
Surdo-mudo	61.819	0,4%	57.537	0,7%	119.356	0,5%
<b>Pessoas com 2 ou mais anos</b>	<b>14.979.335</b>	<b>63,1%</b>	<b>8.760.636</b>	<b>36,9%</b>	<b>23.739.971</b>	<b>100%</b>

Fonte: Inverno (2008, p. 89) adaptado de INE (2016)

Na seção que segue, o foco é dado nos aspectos sócio-históricos de Angola, em um contexto forte de dominação no território africano pelos portugueses, desde o período colonial até um pouco antes da formação de uma nação independente do domínio português, havendo assim, uma independência tardia.

### 3.2.1 História social de Angola

A República de Angola é um país que se localiza na costa ocidental do continente africano e possui uma população de cerca de 26 milhões de pessoas. Depois do Brasil, Angola se destaca na segunda posição em termos de extensão territorial e cada vez mais a língua portuguesa tem se tornado majoritária em seu território, mesmo com a presença de outras línguas com estatuto de língua nacional, como o quimbundo, o quicongo e o umbundo, línguas africanas do grupo bantu.

As relações portuguesas com a África são antigas. O primeiro contato data de 1415 com as sociedades subsaarianas. Em 1448, é a vez da Costa da Mauritània e daí então, vão sendo estabelecidos relações de comerciais entre eles. Em relação à Angola, foco de discussão desse estudo, só acontece quando o português Diogo Cão, por volta de 1482 a 1484 chega ao Reino do Congo, na parte norte de Angola, onde viviam os povos da língua banto. Nesse primeiro momento da chegada dos portugueses, são estabelecidas relações diplomáticas e há um intenso contato já havendo influências na religião e na língua. Já no final da década, por ordem do rei D. João II, os congueses presenciam a primeira tentativa de Portugal de inserir a língua portuguesa no Reino sem falar nos artigos de fé, dos batismos e de aprendizado dos mandamentos. Semelhantemente ao ocorrido em terras brasileiras, os portugueses em Angola, por meio de justificativas ideológicas faziam a conversão dos angolanos ao cristianismo, eram batizados e recebiam assim a doutrina cristã.

Só por volta de 1570, é que os portugueses liderados por Diogo Cão e, após conquistarem a confiança das lideranças políticas, começam a implementar o projeto colonizador de Portugal. Inicia-se assim, a consolidação do território angolano com a chegada de famílias camponesas da Europa composta por 100 famílias e 400 soldados portugueses, muitos deles degredados e condenados em Portugal.

Nos séculos XVII e XVIII, o poderio português é bastante evidente e concretizado pela formação de entrepostos comerciais a fim de abastecer o comércio escravocrata enviando africanos em regime escravocrata para as plantações de cana de açúcar no Brasil, principalmente, Bahia e Pernambuco. Segundo Caregnato (2010),

os portugueses desenvolveram um sistema organizado, isto é, as lideranças políticas eram cooptadas e auxiliavam o poderio português na apreensão de indivíduos, que eram destinados aos portos principais de Luanda e Benguela e, posteriormente, enviados principalmente ao Brasil, colônia portuguesa na América, que necessitava de mão de obra para a manutenção do sistema escravocrata. (Caregnato, 2010, p. 08)

Desse modo, é evidente constatar que um dos maiores interesses e movimentos dos portugueses nas terras africanas acontecia pelo fato do comércio lucrativo de negros escravizados. É evidente observar que o grande interesse dos portugueses, nas terras africanas, decorria do rentável comércio de negros africanos, o tráfico, o que serve de destaque para um dos maiores processos migratórios já presenciados e realizados por vias marítimas. Vale destacar ainda que, segundo Fage (2010, p. 305) os

[...] interesses dos Portugueses na África ocidental banto, em Angola, não estavam ainda totalmente definidos antes 1680, e só nessa data é que outros comerciantes europeus interessados na aquisição de escravos para as Américas começaram a dedicar muita atenção ao litoral entre Angola e os Camarões. (Fage, 2010, p. 305)

No século XIX, Luanda e Benguela já tinham se tornado os maiores centros sociais e econômicos de Angola e, com a abolição da escravatura em 1858, o decreto exigia que todas as colônias de exploração portuguesas cessassem o seu regime colonialista de escravidão. Diante dessa proibição, Portugal estimula cada vez mais a ida de portugueses para as colônias e parte para “civilizar” essas colônias. Essa história de civilização como espécie de dominação de povos já é conhecida na história dos portugueses. Segundo Mingas (2000), o processo de colonização de Angola não tinha apenas uma característica de exploração, mas também de povoamento, pois, no sistema de substituição, uma parcela desses colonos expatriados misturava-se em meio aos povos autóctones e geravam uma porcentagem muito elevada de mestiços.

Durante o período colonial, a sociedade angolana estava dividida basicamente em dois grupos: os portugueses e os angolanos. Os primeiros eram agentes da metrópole portuguesa e seus colonos eram os instrumentos de colonização. Ainda conforme a autora, com o tempo havia um número maior de portugueses nascidos em terras angolanas do que os chegados da Europa. Nesse processo de colonização, já se observava uma separação, uma discriminação na sociedade pois os colonialistas podiam ser os de primeira, eram brancos e isso os conferia privilégios de modo que podiam participar de altos cargos de chefia e ajuda econômica, enquanto os de segunda, os nascidos em Angola, mestiços, sem direitos de participação. Assim, à medida que a população de portugueses nascidos em Angola ia superando, aos poucos, a quantidade de portugueses nascidos na Europa, é que começaram a fazer distinções entre os brancos de primeira e de segunda.

Sobre o grupo dos angolanos, encontravam-se divididos em os assimilados e os indígenas. Os assimilados eram subdivididos na sociedade colonial em passivos e ativos. Para os filhos

dos assimilados era permitido que frequentassem as escolas junto com as crianças portuguesas. Os primeiros, eram considerados alienados, pois não assumiam a sua cultura, seu povo e estavam indiferentes à sua condição de seres humanos. O discurso colonial é empregado para diminuir a representatividade do colonizado como indivíduo. Esses seres, agora submissos, eram apenas tipos degenerados que precisavam de disciplina. Deste modo, a supressão da sua identidade e a busca por uma evolução civilizatória justificavam a conquista e qualquer tipo de ação autoritária. Mesmo a identificação com o passado precisava ser rompida, para que ele pudesse buscar uma identidade a ser admirada. Para Fanon (2008),

Todo povo colonizado — sua originalidade cultural — toma posição diante da linguagem da nação civilizadora, isto é, da cultura metropolitana. Quanto mais assimilar os valores culturais da metrópole, mais o colonizado escapará da sua selva. Quanto mais ele rejeitar sua negridão, seu mato, mais branco será. (Fanon, 2008, p. 34)

Em relação aos assimilados ativos, considerados “negros diferentes” ou de “pele negra e alma branca” pelos colonialistas se distinguiam bastante dos passivos, pois estes tinham o ensejo de libertar o povo e o país da dominação portuguesa. Estes tinham a tenacidade de perceber dentro de um sistema que foi imposto pela exploração e opressão de uma classe que domina, “eram aqueles que se apropriaram de tudo o que lhes foi possível aprender, [...] utilizaram a língua do colonizador para denunciar o colonialismo português através dos jornais [...] da poesia militante e revolucionária, [...] por intermédio de actividades políticas”, (Mingas, 2000, p. 46). Quanto aos indígenas<sup>34</sup>, aqueles que pagavam impostos, a condição para se tornarem assimilados era aprendendo a falar corretamente e adotando os costumes portugueses. Segundo Lucchesi (2015, p. 27), “A universalização das representações que atendem a interesses particulares, e não universais, está na base da construção da hegemonia ideológica da classe dominante”. Nesse aspecto, em termos linguísticos, a história do Brasil se repete. A imposição da língua portuguesa acaba sendo condição *sine qua non* para se fazer parte de uma sociedade preconceituosa e racista em que o falar considerado inculto pela classe dominada, ou variedades nativas, não tem vez para aqueles que se sentem superiores.

Conforme Mingas (2000, p. 49), cada grupo pertencente à sociedade angolana equivalia a um grupo linguístico, os portugueses com o monolinguismo minoritário (português), os assimilados com o bilinguismo (kimbundo/português) e os indígenas com o monolinguismo majoritário (kimbundo). Importante destacar sobre a parcela da população que era bilíngue, ou

---

<sup>34</sup> Indivíduos de raça negra ou seus descendentes que nasceram ou viveram nas terras, mas não tinham hábitos e costumes portugueses.

seja, falava o português como segunda língua, o “pretoquês”, um português morfo-sintaticamente modificado. Sobre isso, a autora afirma que essa língua se tornou

[...] elemento de ridicularização em peças de teatro, levou a que os ‘Assimilados’, na quase sua totalidade, fossem os primeiros a contribuir para a concretização da política linguística dos colonizadores: *eles evitavam falar a sua língua materna e proibiam mesmo os seus filhos de o fazer*. (Mingas, 2000, p. 50. Destaque da autora.).

O fato de essa estratégia ser utilizada como garantia para a imposição da língua portuguesa como língua oficial, de alguma forma criou um terreno fértil para o surgimento de novas pessoas que falassem o português como L1, mesmo em contexto de pais bilíngues. Outro fator que contribuiu também para a concretização da LP foi a própria taxa de analfabetismo do povo angolano resultando em uma transmissão para as novas gerações de uma variedade de português com interferências das línguas locais, em especial do quimbundo.

Para se ascender socialmente era importante ter o conhecimento da língua portuguesa de modo que, saber falar o “bom português” era falar sem interferências de suas línguas nativas, por isso, de acordo com as leis promulgadas, nem nas escolas católicas, nas mídias (jornais, panfletos) e nem nos contatos sociais as línguas indígenas eram permitidas. Mesmo assim, poucos foram os angolanos que conseguiram a ascensão em relação aos estudos em uma universidade, fato que ajudou a uma identidade linguística portuguesa com interferências, principalmente da língua quimbundo. Vale ressaltar também que a falta de gratuidade, a não obrigação e permissão por grande parte dos angolanos de frequentarem a escola, favoreceu o uso das línguas nacionais, ou seja, as situações de multilinguismo ainda persistiam mesmo com a proibição.

Após a II Guerra Mundial, em meados do século XX, o colonialismo europeu é enfraquecido e acaba fortalecendo os movimentos que exigiam a libertação do domínio português e desejavam a independência do país. Em 1975, Angola torna-se independente com a queda do regime português ditatorial de Salazar, e nesse novo contexto sócio-histórico e político, a Língua Portuguesa é de fato reconhecida como língua oficial da ex-colônia portuguesa. Nesse momento, já existe uma preocupação para a implantação de políticas linguísticas no país dando abertura para a valorização, a princípio, das seis línguas nacionais de diferentes grupos étnicos, em sua maioria dos grupos banto, como o umbundo, quimbundo, kikongo, cokwe, kwanyama e nganguela.

Os principais grupos, o Movimento Popular de Libertação de Angola (doravante MPLA), de ideologia marxista, União Soviética (URSS), a Frente Nacional para a libertação de Angola

(FNLA), financiada pelos EUA e a União Nacional para a Independência Total de Angola (UNITA), apoiada pelo regime da África do Sul, *Apartheid*, embora tivessem objetivos em comum, tinham ideologias distintas o que provocou uma longa Guerra Civil que só veio terminar no início deste século, em 2002. Com os violentos efeitos da Guerra Civil para os angolanos, muitos deles acabaram migrando das áreas mais rurais para as regiões costeiras em busca de melhores condições de vida e, nesse contexto, tiveram que aprender o português como L2 cujos descendentes passariam a utilizar o português como L1.

### 3.2.2 Formação do Português Angolano

Depois que Angola conquistou a sua independência, uma nova fase foi instaurada no país na tentativa de pôr em prática, em um país livre, as vontades de uma nação que ansiava, há muitos anos, a liberdade das amarras amargas do domínio europeu e das intensas guerras civis. Diante disso, o artigo 1º da Constituição de Angola (1975) propõe:

Total libertação do povo angolano dos vestígios do colonialismo e da dominação e agressão do imperialismo e a construção dum país próspero e democrático, completamente livre de qualquer forma de exploração do homem pelo homem, materializando as aspirações das massas populares. (Art. 1º, Lei Constitucional, 1975)

Zau (2011) reitera que a língua portuguesa, em território angolano, não pode ser considerado algo alheio à história do colonialismo, muito pelo contrário, ela é uma clássica representante desse vestígio colonial, no entanto, segundo o autor (Zau, 2005, p. 112 -113), “não restam dúvidas quanto à sua utilidade para a formatação de uma nova identidade nacional à custa de ameaça da diversidade”. Desse modo, a língua portuguesa, mesmo ainda estando ligada a um passado cruel de imposição, enquanto uma ferramenta para a comunicação nacional, contribuiu para tal formação de uma nova identidade. Assim, “[...] a angolanidade de que é, paradoxalmente, veículo e cofundadora, ganha com a língua portuguesa e esta com aquela”. (Zau, 2011, p. 112-113)

A língua portuguesa, que antes representava opressão, passa a ser considerada um símbolo de unidade nacional levando a cada cidadão africano um sentimento de reafirmação de sua angolanidade, atrelando as suas raízes históricas com um país, agora, independente. Sendo assim, a língua portuguesa é vista como um elemento sociocultural que uniria o passado extremamente marcado nas histórias vividas pelos angolanos a um futuro, a uma esperança que se renovava a partir da libertação do seu povo. Zau (2011, p. 113-114), afirma que “o conflito

armado provocou simultaneamente a desterritorialização da população e a territorialização da língua portuguesa” e ainda destaca três contribuições fundamentais para a ascensão da última:

- a) A guerra forçou à convivência de pessoas de etnias diferentes nos mesmos espaços, na medida em que a maioria da população do interior abandonou as suas zonas de origem em busca dos grandes centros que oferecessem maior segurança ameaçada pela guerra;
- b) A partilha de espaço comum acima referida ajudou a encurtar as fronteiras identitárias, a diminuir querelas e a inculcar a ideia de “um só povo, uma só nação(!)”;
- c) A língua portuguesa foi um instrumento fundamental em todo este processo, funcionando como elo de ligação, ou seja, com funções de língua franca, por ser aquela através da qual populações de etnias diferentes podiam comunicar entre si, em todas as suas necessidades. (Zau, 2011, p. 113- 114)

Vale ainda enfatizar, com base no autor supracitado, que a língua portuguesa tem se destacado de uma maneira oficial no mercado e nas transações por alguns motivos elencados a seguir:

- i) no quadro das línguas de Angola, a língua com maior tradição académica (mais bem estruturada e estudada) e com maior projecção nacional e internacional;
- ii) é a única língua com condições estruturais para assegurar a escolarização em todo o território nacional;
- iii) é a única língua isenta de conotação étnica, aspecto fundamental para evitar o risco da ameaça de “guerra linguística” interétnica;
- iv) é a única língua que reúne condições de ser língua da diplomacia, ou seja, 115 línguas não só das cimeiras<sup>35</sup> entre os PALOP (Países Africanos de Língua Oficial Portuguesa), como também entre cada um deles e outros países da CPLP (Comunidade dos Países de Língua Portuguesa), ou com outras instituições internacionais. (Zau, 2011, p. 114-115)

Interessante reafirmar, portanto, que, durante todo o período colonial, a situação das línguas nacionais de origem africana foi fortemente combatida e, de maneira depreciativa, consideradas “dialetos”. É de se pensar que, quando se define determinada língua como oficial em detrimento de outras, o carácter eurocentrista de um grupo hegemónico ainda se faz presente. No caso da língua portuguesa no contexto educacional de Angola, o uso nos ambientes escolares e administrativos não privilegia a realidade sociolinguística do país pois, visto que a língua portuguesa reflete “[...] ideologias, culturas e modos de ver o mundo um tanto distantes da realidade pluridiscursiva, plurilíngues e diversificada africana” (Bernardo; Severo; 2018, p. 213), o que agrava ainda mais se se pensa nas áreas interioranas, rurais do país que são mais constitutivas de falantes das diferentes línguas nacionais. Assim,

[...] a língua portuguesa como a língua oficial de comunicação nacional, hegemонizando-a e tornando-a um instrumento de manutenção ideológica do monolinguismo. Esse monolinguismo arraigado à língua portuguesa é

<sup>35</sup> Palavra utilizada em Portugal que equivale à cúpula, congresso, conferência, reunião de alto nível.

sustentado por uma retórica que localiza as demais línguas angolanas sob o escopo de “línguas nacionais” (Bernardo; Severo; 2018, p. 213).

Inverno (2018) faz uma crítica ao discurso dominante nos PALOP sobre a homogeneidade secular da língua portuguesa no território africano e afirma que, quando o português se tornou a língua oficial no pós independência como em Angola, Moçambique e, em menor grau, de Cabo Verde, Guiné-Bissau e São Tomé e Príncipe, não foi um processo generalizado. Desse modo, ela tenta desmistificar a difusão do monolinguismo nesses países e aponta para a difusão de variedades nacionais influenciadas por línguas diversas do grupo banto, como ocorre em Angola. Em um país que ainda vive intensamente os contatos linguísticos de diferentes etnias e grupos, a língua portuguesa se configura um ensino excludente que não reflete o cenário multilinguístico e que ainda gera tensão entre os povos.

### **3.2.3. Relações oralidade/escrita nas culturas africanas**

Discutiu-se sobre a influência da chegada dos portugueses nos modos de pensar dos angolanos e alguns fatos da sócio-história de Angola nesta pesquisa. Nesta subseção, convém refletir sobre o processo tardio da língua escrita na comunicação da humanidade e, que mesmo assim, ainda hoje se percebe uma certa superioridade da escrita em relação à fala. Destacam-se, aqui, os povos africanos como povos essencialmente de tradições orais.

A tradição oral e a tradição escrita são definidas para Calvet (2011) como formas de comunicação linguística que podem estar presentes em uma sociedade. A partir de um leque de possibilidades, o autor afirma que a maioria das sociedades europeias atuais estão inseridas, conforme sua simplificação tipológica, em sociedades de tradição de escrita antiga devido ao fato de a língua escrita ser “[...] aquela que se utiliza na comunicação oral cotidiana” (Calvet, 2011, p. 11) enquanto os países que foram colônias de Portugal, como é o caso de alguns países africanos, a exemplo de Angola e, do Brasil, na América Latina, são sociedades que tiveram um processo de prática escrita bem recente, visto que uma língua diferente daquelas faladas localmente foi introduzida pelo colonialismo português. Essa herança imposta aos povos africanos, povos de forte tradição oral, deixa evidente que a escrita, certamente, serviu como meio de exploração, uma das formas de poder e dominação. Nesse contexto, pode-se dizer que o processo de alfabetização nas terras colonizadas se torna ainda mais difícil, pois, antes de ensinar essa nova língua, há antes de tudo um processo de criação de uma escrita.

Conforme Marcuschi (2010, p. 17),

[...] todos os povos, indistintamente, têm ou tiveram uma tradição oral, mas relativamente poucos tiveram ou têm uma tradição escrita, isto não torna a oralidade mais importante ou prestigiosa que a escrita [...] os usos da escrita, no entanto, quando arraigados numa dada sociedade, impõem-se com uma violência inusitada e adquirem um valor social até superior à oralidade. (Marcuschi, 2010, p. 17).

Por muito tempo, segundo Fonseca (2016), as questões de oralidade ficaram relegadas ao segundo plano, pois era forte a defesa da supremacia da escrita e a crença de achar que a oralidade era uma modalidade primitiva da língua, um estágio de comunicação humana anterior à escrita. Ainda, segundo a autora, houve a desconsideração das produções culturais orais da memória, das tradições africanas, sendo valorizadas apenas aquelas produções escritas e conservadas em bibliotecas. A oralidade, as tradições orais são importantíssimas para a valorização das práticas que possibilitam o ouvir e o repetir que é passado de geração a geração. Nesse contexto, pode-se entender não somente a variedade oral utilizada pelos falantes, mas também como um legado de textos orais próprios de uma determinada cultura.

Se para muitos povos africanos a palavra é sagrada porque é intermediada pela força do Ser Supremo, para outros povos do continente e para várias culturas que tiveram contato com os saberes africanos através da escravidão, a fala guarda a energia da força vital que está presente no sopro que deu vida ao homem e naquilo que ele tira de si através da palavra proferida. De alguma forma, os estudiosos percebem a oralidade como uma manifestação inerente ao homem, permitindo-lhe expressar a integridade de seus pensamentos (Fonseca, 2016, p. 13).

No contexto de Angola e em outras culturas que passaram por histórias parecidas, enquanto se tem uma língua escrita com características marcadamente europeia por causa da imposição da LP, a oralidade ainda é bastante africana (Torres; Araujo, 2024, p. 128). Esse aspecto vem fortalecer bastante a bipolaridade entre o oral e o escrito, pois não leva em conta a existência da escrita antes da chegada dos portugueses no período colonial pelos povos africanos. Hoje, já há uma tendência de alguns escritores e intelectuais, a exemplo de Luandino Vieira e Agostinho Neto, em representarem nas suas produções os costumes africanos; há o uso da linguagem intencional de suas origens africanas com representações da língua falada pela população mais pobre, das regiões suburbanas de Luanda, os chamados musseques, como da própria morfossintaxe e léxico, especialmente do Quimundo, configurando assim, uma espécie de mestiçagem e mesclagem de culturas orais e escritas. Conforme Leite (1998, p. 17), a oralidade que predomina no território africano não resulta de uma natureza africana, mas sim

“[...] de condições materiais e históricas [...] muitos críticos partem do princípio de que há algo ontologicamente oral em África, e que a escrita é um acontecimento disjuntivo e alienígena para os africanos.”

Sobre esse aspecto, as sociedades que têm tradições orais possuem suas próprias especificidades e deveriam elas mesmas decidirem a maneira como utilizarem a escrita em certo momento ou não, conforme as suas necessidades. É humana a necessidade dos distintos povos se expressarem por meio das linguagens e, muitas vezes, a(s) língua(s) utilizada(s) por eles podem se manifestar por meio dos diversos gêneros orais, escritos e multimodais. Em algum ponto, acredita-se que é possível que outras necessidades surjam nas comunidades de tradições mais orais e o uso escrita passaria a se relacionar às outras diversas condições de produção e sentido para essas sociedades. Incontestável é o fato de que a língua se molda aos diversos usos que os cidadãos fazem dela na sociedade. Calvet (2011, p. 145) reitera que

[...] a escrita não é uma simples transcrição da língua, ela tem implicações sociais muito mais profundas, exatamente como a oralidade, que não é a ausência da escrita. Por isso não se deveria impor arbitrariamente a sociedade oral o código escrito que não nasceu dela, não se deveria prolongar a dominação de uma sociedade sobre outra impondo-lhe esse código em um momento histórico e em condições que ela não escolheu. (Calvet, 2011, p. 145)

Na subseção seguinte, a discussão estará voltada para o plano educacional, de políticas linguísticas implantadas, sem perder de vista o cenário multilíngue angolano, os efeitos perversos gerados pela necessidade de saber falar bem o português, de adquirir hábitos europeizados, de conquistarem a condição de assimilado e de ser um cidadão português.

### **3.2.4 Multilinguismo em Angola: línguas locais e língua portuguesa**

Muitos aspectos foram motivadores para a expansão dos usos da escrita nas sociedades tecnológicas como a emergência de estados nacionais, as mudanças socioeconômicas, o desenvolvimento das ciências, e até mesmo a dominação, padronização e imposição de uma variante linguística na escola e no meio administrativo das cidades, por exemplo. Nesse contexto, permite-se afirmar que a situação em terras angolanas pode ter uma estreita relação com a ideia de civilização, progresso ou modernização dividindo assim grupos orais e letrados. Essas diferentes impulsões para a ampliação da escrita, em Angola, não aconteceriam diferentemente de outras culturas, no entanto, é necessário refletir como são e quais os efeitos

dessas práticas de letramento para grupos minoritários ou para aqueles considerados subjugados pelos grupos dominantes.

A Igreja Católica, por exemplo, em suas missões, foi uma fiel representante do reinado português e uma das grandes encarregadas pela conversão dos povos locais angolanos e pelo ensino, o que contribuiu fortemente para que as diferentes comunidades étnicas não se inscrevessem num espaço que favorecesse a discussão e valorização do lugar do outro e restringindo a participação dos locais a produzirem sentidos e novas formações discursivas. Sendo assim, os sujeitos angolanos não podiam compartilhar suas próprias experiências, seus sentimentos e conhecimentos de mundo.

Um outro fator que também acabou contribuindo para a não valorização das línguas locais pelos missionários portugueses no processo de ensino em Angola foram as próprias línguas africanas, que por serem basicamente línguas orais, foi crucial para que o modelo de missionarização utilizasse a língua portuguesa com o intuito de converter e dominar os nativos de acordo suas ideologias (Makoni, 1998).

Segundo Oliveira (2018, p. 241),

[...] tal modelo visava a adaptação de línguas orais às línguas escritas para transpor aos africanos o gênero religioso em suas línguas, realizando, portanto, o método conversional. No entanto, este processo implica uma outra linguagem que não é de cunho epilinguístico, em sua forma oral, mas uma língua em seu estado objetificado, metalinguístico. (Oliveira, 2018, p. 241).

O início do processo de silenciamento e isolamento das línguas africanas tem afetado até os angolanos atualmente. Famílias africanas preferem que seus filhos não aprendam as línguas locais pois, “[...] muitas famílias receiam que o tempo necessário para aprender a segunda língua, nacional, influencie as restantes atividades dos mais novos, optando apenas pelo idioma oficial” (Tchimboto, 2016, s.p)<sup>36</sup>. Para muitos angolanos, com a hierarquização da língua portuguesa em seu território, a aceitação pela variedade nacional é a que deve ser aprendida e as situações de bilinguismo/multilinguismo são consideradas um defeito. O autor ainda aponta que a memória dos locais, falantes de Umbundo, é forte ao relatar da lembrança da placa de madeira que os chamavam de “burros” ao chegarem no ambiente escolar.

Como as línguas locais podiam representar ameaça ao governo português, houve promulgação de leis e artigos como as seguintes, a Lei nº 1031, de 6 de março de 1921.

---

<sup>36</sup> <http://ipol.org.br/mais-de-sete-milhoes-de-angolanos-falam-linguas-nacionais/>

... Artigo 2.º ... nas escolas católicas, é proibido ensinar as línguas indígenas;

... Artigo 3.º ... a utilização das línguas indígenas no catecismo não é permitida a não ser como auxiliar durante o período de ensino elementar da língua portuguesa;

...  
 § 1. É proibido o emprego das línguas indígenas ou qualquer outra língua, à excepção do português, por escrito ou por panfleto, jornal, ... na catequese das missões, nas escolas e em todas os contactos com as populações locais... (Mingas, 2000, p. 32-33).

A repressão era tão intensa que, nas escolas, por exemplo, em casos que um filho de assimilado resolvesse se expressar, falar na sua língua local, o quimbundo, era muito grande a chance de seus pais terem problemas com a polícia. A PIDE<sup>37</sup>, a polícia portuguesa, coibia as atitudes consideradas subversivas, inclusive em relação às próprias vestimentas.

Os angolanos se sentiam inferiorizados, ridicularizados e com vergonha por terem como língua materna, suas línguas nacionais, desprestigiadas e, por um longo período, eles lutaram para não falar suas línguas locais. A variante portuguesa utilizada por eles, um português que nascia com interferências de suas línguas nativas, em especial, o quimbundo, era pejorativamente chamada de “pretuguês”, “português de preto”, o que segundo Mingas (2000), contribuiu bastante para a concretização de uma política linguística que aniquila a língua dos povos autóctones e privilegia o estrangeiro. Essa situação demonstra uma política linguística que contribui para o racismo linguístico visto que a categorização e a avaliação feitas evidenciavam as piores qualidades e geravam também consequências sociais e em questões identitárias dos povos desprezados. Conforme Foucault,

[...] quando penso na mecânica do poder, penso em sua forma capilar de existir, no ponto em que o poder encontra o nível dos indivíduos, atinge seus corpos, vem se inserir em seus gestos, suas atitudes, seus discursos, sua aprendizagem, sua vida quotidiana. (Foucault, 2016, p. 215).

Pode-se pontuar ainda que o Estatuto do Indigenato, decreto imposto pela colônia, legitimou a subjugação do povo angolano por ser considerada uma política que exigia uma normatização, padronização daquilo que era diverso. Com o enaltecimento da língua portuguesa, houve não só a desvalorização das línguas nacionais, mas também da desvalorização dos sujeitos. Tal cenário só apresentou outros olhares na década de 60 quando, devido às pressões africanas, já se pensava em criar meios para se discutir questões ligadas à

---

<sup>37</sup> A Polícia Internacional e de Defesa do Estado (PIDE) era a polícia portuguesa que atuou entre os anos de 1945 e 1969 cujo papel era repreender todas as formas de oposição ao regime político no auge do Estado Novo.

resistência diante das políticas segregacionistas, preconceituosas, com enfoque no multiculturalismo, na valorização das línguas e nas identidades nacionais.

A língua portuguesa não perdeu sua supremacia e continuou sendo a língua oficial da República de Angola mesmo após a independência do país, conforme reitera o Artigo 19º da Constituição (2010, p. 9) ao tratar das línguas “1. A língua oficial da República de Angola é o português. 2. O Estado valoriza e promove o estudo, o ensino e a utilização das demais línguas de Angola, bem como das principais línguas de comunicação internacional”. Mais adiante, no artigo 21º do mesmo documento (2010, p. 11), a alínea *n* afirma que é tarefa fundamental do Estado “proteger, valorizar e dignificar as línguas angolanas de origem africana, como património cultural, e promover o seu desenvolvimento, como línguas de identidade nacional e de comunicação”. A Constituição de Angola no Artigo 87º (2010, p. 30) que trata do Patrimônio histórico, cultura e artístico assegura que:

1. Os cidadãos e as comunidades têm direito ao respeito, valorização e preservação da sua identidade cultural, linguística e artística.
2. O Estado promove e estimula a conservação e valorização do património histórico, cultural e artístico do povo angolano. (A Constituição de Angola no Artigo 87º, 2010, p. 30)

Considerando o que rege a Constituição angolana, não restam dúvidas que muitas mudanças ocorreram. Aos poucos, a valorização de línguas nacionais vai delineando-se e, que são tão importantes para essa nação tão multilíngue e de culturas tão plurais. Assim, o “[...] estado promove e assegura condições humanas científico-técnicas, materiais e financeiras para a expansão e generalização da utilização e do ensino de línguas nacionais” (Angola, 2001, p. 5). Os povos locais angolanos pagaram um alto preço em todo o percurso para apenas, em um pouco mais de 10 anos, o ensino das principais línguas nacionais se tornarem realidade nas escolas nas 18 províncias de Angola, embora as realidades de cada província ainda sejam muito díspares e andem em ritmos desiguais. Quando as línguas endógenas são colocadas em evidência em comunidades plurilíngues, os indivíduos se sentem mais valorizados, passam a conhecer melhor a sua própria cultura e a valorizar ainda mais a sua história, seus antepassados e suas diversas manifestações da linguagem, incluindo a língua falada. Há um ganho muito grande da população com a aprendizagem e o uso das línguas angolanas nas escolas, principalmente, na desconstrução sobre quem são, a representatividade e importância das suas culturas no território africano que promovam o respeito da diversidade cultural e linguística.

Na subseção seguinte, pretende-se discutir um pouco mais sobre o ensino da língua portuguesa hoje em Angola relacionando com as variedades endógenas.

### **3.2.5 Ensino da língua portuguesa hoje em Angola e as variedades endógenas**

“o povo ainda não é vencedor, continua vencido: não há escolas para todos, e a escola que existe é antes contra o povo que para o povo” (Soares, 1994, p. 09)

Embora muitos angolanos acreditem que exista um reconhecimento, uma valorização para as línguas nacionais pelo governo angolano depois da efetiva independência, muitas famílias ainda possuem os valores que, por muito tempo, foram silenciados, assim, até hoje muitos carregam uma carga negativa ao tratar do ensino da língua nacional ser considerado uma perda de tempo nas escolas e até no uso no seio familiar. Segundo Silva, Sene e Araújo (2018):

Compreender as crenças é uma importante ferramenta que pode revelar o valor social de uma língua e como ela é usada para refletir, por exemplo, “símbolo de pertencimento a determinado grupo”, isto é, uma identificação na sociedade. [...] as crenças que um falante tem de sua própria variedade/língua, se a considera apropriada ou não perante outras variedades/línguas, fará com que ele a utilize e, de certa forma, propague sua maneira de falar. (Silva; Sene; Araújo, 2018, p. 341).

De acordo com as sistematizações das respostas da pesquisa realizada por Silva *et. al.* (2018), os resultados apontam que a maioria dos informantes da pesquisa reconhecem que a língua nativa é a mais “bonita” que o português e que as crianças devem aprender as línguas nacionais na escola, no entanto, há resultado de uma informante que ignora o fato de que manter a sua língua materna é manter sua cultura viva, um clássico exemplo de rejeição linguística de sua língua falada, visto que com quase 40 línguas maternas no território, segundo ela, teria um processo complicado para se ensinar. Há de se destacar aqui que as línguas de origem são mais faladas pelos angolanos nas áreas suburbana e rural utilizando-as para rituais de passagem, casamento ou de evocação dos antepassados. No entanto, com o abandono de angolanos das áreas rurais para as grandes cidades, muitos deixaram de usar as suas línguas de origem e passaram a usar o português como língua franca. Ndombele e Timbane (2020) afirmam que o português falado nas cidades adquiriu características próprias, o português angolano, enquanto a promoção e valorização do ensino para as línguas locais foi algo mais presente na teoria, no “papel”, já que, com a chegada da dita educação “moderna”, “civilizada”, “europeia”, mudanças aconteceram, como é o caso de uma educação mais baseada na idade para uma mais

escrita, visto que muitos eventos de letramento acabam tendo mais valor quando é escrito do que oralizado.

Ainda conforme os autores supracitados, os angolanos

[...] sempre tiveram uma educação tradicional. Essa educação ocorreu nas comunidades e se relacionou à cultura que molda as regras de ser e de estar em sociedade. Com a colonização nos foi imposta uma educação moderna, de viés europeu que nos oferece oportunidades e vantagens econômicas. O modelo, a estrutura e os objetivos da educação tradicional visavam perpetuar os hábitos e costumes dos grupos étnicos. A educação tradicional dos povos bantu e khoisan em Angola priorizava a oralidade e as línguas oficiais dessa educação eram as línguas autóctones. A ligação entre a teoria e a prática era muito nítida porque era uma educação para a vida prática do cotidiano (Ndombele; Timbane, 2020, p. 292).

Numa sociedade que privilegia a escrita na administração, na imprensa e em várias outras situações mais formais do cotidiano, as crianças em Angola, quando iniciam seus estudos, geralmente de forma tardia, encontram muitas dificuldades. Precisam ser alfabetizadas, aprender a ler e escrever e esse processo é árduo pois necessitam para fazer parte dessa sociedade, aprender uma língua “inventada” que não reflete suas realidades comunicativas. Muitos professores carecem de formação e aperfeiçoamento para lidar com os alunos oriundos de etnias e línguas locais distintas. É comum que as crianças cheguem às escolas falando mais de uma língua, têm dificuldades de entender as aulas que são ministradas em língua portuguesa orientados por manuais e materiais de base ainda europeia. Como não existem materiais adequados a esse novo contexto, os alunos passam a aprender o português como língua materna, por isso, sentem muitas dificuldades com produções orais e escritas.

Língua e cultura andam juntas e escola nenhuma deveria desconsiderar essa questão que é tão complexa e fundamental. Conforme Kleiman (1995), os fenômenos de letramento vão muito além desse mundo preso à escrita que as instituições ditam visto que os letramentos são considerados práticas discursivas de diferentes grupos sociais que, não necessariamente, se referem especificamente ao ato de ler ou escrever. Ao diferenciar letramento de alfabetização, Marcuschi (2010) afirma que o primeiro é um processo de aprendizagem em contextos informais, se relaciona com o social e histórico da leitura, ou seja, o que as crianças levam para a escola partindo de suas vivências, suas relações com seus familiares e amigos de diferentes contextos de letramento, por necessidades ou motivações que lhes são próprias; o segundo, já está atrelado aos contextos formais, e responde ao domínio e sistematicidade das habilidades de ler e escrever como é o caso das instituições escolares. Deste modo, quando as crianças angolanas adentram os espaços escolares, levam seus discursos, suas histórias, suas memórias,

mas, muitas vezes, acabam apenas racionalizando a repetição formal, negam sua autoria, sua identidade, o que não é exclusivo dos alunos, professores também podem vivenciar esse silenciamento.

O processo de escolarização das diferentes comunidades étnicas angolanas se encontra ligado à decodificação do código linguístico e se baseia majoritariamente em textos escritos, sendo que todo o processo de formação e o progresso é feito pela escrita. Nesse modelo autônomo de letramento, há o respaldo do modo tecnicista de interpretar a língua cujo processo de padronização da escrita ocorre pela divulgação ampla de dicionários e gramática. Existe assim, a anulação da vida do sujeito visto que o ensino é realizado por meio da linguagem, mas de maneira descontextualizada. Ressalta-se, ainda, que a política desempenhada no contexto colonial de Angola fortaleceu a construção de um imaginário de língua que não deveria refletir a realidade linguística dos grupos étnicos, há uma característica engessada das línguas.

A escrita deveria ser considerada um dos meios, tecnologia, um sistema simbólico a ser utilizado em contextos específicos e com objetivos específicos. (Scribner; Cole, 1981). E, quando se fala no ensino de línguas, os próprios currículos e práticas escolares, embora considerem importantes trabalhar com as práticas orais, a escrita não perde a supremacia, na prática, é como se fosse a única modalidade digna de atenção. Tfouni (1995, p. 19) afirma que “A relação entre a escrita e a oralidade não é uma relação de dependência da primeira à segunda, mas é antes uma relação de interdependência, isto é, ambos os sistemas de representação influenciam-se igualmente”. Diante disso, entende-se que a escola perde de vista que oralidade e escrita devem ser atividades de comunicação e de práticas sociais situadas para que seja possível promover usos reais da língua (Marcuschi, 2010).

Até o que foi discutido aqui, pode-se perceber que o sistema colonial deixou uma forte herança impregnada na vida dos angolanos até os dias de hoje. Língua e discursos serviram como grandes instrumentos de dominação de povos pluriétnicos e multilíngues. As vozes, as culturas, as línguas dos povos locais foram silenciadas e, só depois de muita luta, no pós-independência do país, que o governo lança novas medidas, novas políticas linguísticas. O multilinguismo não deveria ser um entrave educacional, mas o discurso ambivalente e contraditório presentes nas práticas, ao mesmo tempo em que reconhece e valoriza a diversidade das línguas locais, prega a orientação é para a língua portuguesa.

Em um país de contatos linguísticos intensos e com diferentes etnias e grupos, a Língua Portuguesa se configura um ensino excludente que não reflete o cenário multilinguístico e que gera tensão entre os povos a ponto de pais negarem as suas raízes, suas histórias, na crença de que seus filhos sejam educados usando a língua imposta, condição para a ascensão social, além

de considerar perda de tempo o ensino das variedades de línguas locais. A língua portuguesa em Angola já não é a mesma que chegou em 1482; muitas transformações ocorreram e, embora muitos escritores tentem resgatar as antigas tradições orais africanas na modalidade escrita, o ensino de língua nas escolas angolanas ainda é europeu, descontextualizado, gramaticizado, fora das situações reais de uso, de produção e comunicação, não promotora dos multiletramentos nem do prestígio das outras línguas faladas em seu território. Em grande parte dos casos, os “dialetos” locais são apenas instrumentos para o ensino da escrita portuguesa.

É fundamental continuar investindo em novas políticas de planejamento linguístico que reflitam efetivamente a diversidade linguística e identitária na formação dos professores, na renovação de práticas escolares mais inclusivas que trabalhem as modalidades oral e escrita (complementares), que trabalhem com o pluralismo de seus estudantes, aproximando-os de contextos mais reais. Escolas que nutram a esperança de um ensino promotor da liberdade das amarras de um passado excludente e preconceituoso, do pensamento crítico, do respeito e do plural são necessárias enquanto aquelas que discriminam, classificam os sujeitos em letrados/iletrados, alfabetizados/analfabetos devem ter suas ações e práticas repensadas.

Na seção 4, a seguir, a discussão é focada em estudos sobre o fenômeno da indeterminação do sujeito, deste as abordagens mais tradicionais, as formas rígidas e limitadas pelas gramáticas normativas, até as estratégias mais inovadoras e que fazem parte do uso real pelos falantes.

## 4 O FENÔMENO LINGUÍSTICO: INDETERMINAÇÃO DO SUJEITO



Fonte: <https://www.jeangalvaocartunista.com/pedag%C3%B3gicas>

Nesta seção, inicialmente, discorre-se sobre as definições dadas sobre as gramáticas tradicionais brasileiras/portuguesas sobre os distintos pontos de vista para definições do sujeito. Análises parecidas ocorrem sobre a indeterminação do sujeito, sem perder de vista as concepções não tradicionais sobre o fenômeno, que apontam outros caminhos, outras estratégias de indeterminação do sujeito pelos falantes que devem ser consideradas legítimas e que já se caracterizam como um fenômeno variável da língua. Desse modo, também são discutidos alguns conceitos do que pode ser considerado determinado ou não, bem como a conceituação entre o que pode ser considerado indeterminação, indefinição e impessoalização. Em seguida, são apresentados trabalhos já realizados sobre o fenômeno da indeterminação do sujeito, e no território brasileiro e no português angolano, e seus resultados são discutidos.

### 4.1 SUJEITO: ALGUMAS CONCEPÇÕES

Antes de iniciar a discussão sobre a indeterminação do sujeito dada pelas Gramáticas Normativas (doravante GTs) e das outras estratégias utilizadas pelos falantes, convém, primeiramente, entender o conceito de sujeito. Antes de qualquer coisa, é necessário entender sobre as diferentes abordagens para a definição de sujeito, para, mais adiante, compreender a visão de indeterminação que se propõe discutir e analisar nesse trabalho. Assim, o primeiro ponto de discussão é sobre o sujeito com base nas definições dadas pelos gramáticos, a exemplo de Almeida (1999); Faraco e Moura (1997), Cunha e Cintra (2016); Cegalla (2008), Rocha Lima (2011), Bechara (2019), entre outros.

Cunha e Cintra (2016, p. 136), declaram que o sujeito é “[...] o ser sobre o qual se faz uma declaração” e que nem sempre o sujeito e o predicado vêm materialmente expressos como em “Andei léguas de sombra dentro em meu pensamento. (F. Pessoa, OP, 59.), já que o sujeito

da ação andar (andei) é indicado apenas pela desinência verbal e é *eu*. Essa visão de sujeito é corroborada por Faraco e Moura (1997, p. 313), em que afirmam que o sujeito é um termo essencial da oração que expressa “[...] o ser a respeito de quem ou de que se faz um a declaração”, como nos exemplos a seguir:

(10) “*A cela da delegacia* jamais foi usada.”

(11) “*Os seus 5 mil habitantes* comem muitas frutas todos os dias.”

Desse modo, o sujeito pode ser representado por um substantivo ou palavra substantivada, por pronome pessoal reto, pronome demonstrativo, relativo, interrogativo ou indefinido, numeral; por oração subordinada substantiva subjetiva e até por pronome pessoal oblíquo. O sujeito pode ser encontrado anteposto ao verbo ou posposto a ele.

Cegalla (2008, p. 324) afirma que o sujeito “é o ser do qual se diz alguma coisa” e, para Bechara (2019, p. 435) “[...] sujeito é uma noção gramatical, e não semântica, isto é, uma referência à realidade designada, como ocorre com as noções de agente e paciente. Assim, o sujeito não é necessariamente o agente do processo designado pelo núcleo verbal”.

## 4.2 A INDETERMINAÇÃO DO SUJEITO

Até aqui, foram apresentados, brevemente, os entendimentos de sujeito dados por algumas gramáticas (Cf. 3.1). Assim, nesta subseção, são discutidas algumas abordagens para o sujeito indeterminado presentes nas gramáticas tradicionais, a exemplo de Bechara (1966, 2019), Cunha e Cintra (2016), Cegalla (2008), Rocha Lima (2011), descritivas para a indeterminação do sujeito, bem como visões de linguistas sobre o fenômeno como: Perini (2002); Neves (2011); Castilho (2019) e Bagno (2011). Assim, além de ser apresentado a perspectiva da GT e descritiva, são apresentados resultados de pesquisas realizadas sobre esse a indeterminação do sujeito, tanto na oralidade quanto na escrita a exemplo de Milanez (1982); Rollemberg et al (1991); Menon; (1994); Setti (1997); Godoy (1999); Santana (2006), Ponte (2008), Teixeira (2008) Carvalho (2010), Teixeira e Almeida (2011), Assunção (2012), Souza, (2014), Mendonça e Nascimento (2015), Dias (2017), Guimarães (2018), entre outros.

### 4.2.1 Abordagens tradicionais para a indeterminação do sujeito

Cunha e Cintra (2016, p. 142-143), consideram as situações de sujeito indeterminado os “[...] casos em que o sujeito não vem expresso na oração nem pode ser identificado”. Desse modo, a indeterminação se daria de duas formas:

**1. Com verbos na 3ª pessoa do plural:**

(12) *Contaram-me*, quando eu era pequenina, a história duns náufragos, como nós.

(13) *Reputavam-no* o maior comilão da cidade.

**2. Com verbos na 3ª pessoa do singular, como o pronome se:**

(14) “Ainda *se viva* num mundo de certezas. *Precisa-se* do carvalho; não se precisa do caniço. *Comia-se* com a boca, com os olhos, com o nariz.”

Conforme os autores supracitados, é ainda possível que ocorra mais de duas maneiras de indeterminação em um mesmo período “Na casa *pisavam* sem sapatos, e *falava-se* baixo”. (Cunha e Cintra, 2016, p. 143).

Cegalla (2008, p. 325- 326) define o sujeito indeterminado “[...] quando não se indica o agente da ação verbal” pois não é possível indicar quem realizou a ação, como por exemplo em:

(15) “Atropelaram uma senhora na esquina.”

(16) “Come-se bem naquele restaurante.”

O autor aconselha as pessoas a não confundirem um sujeito oculto com um indeterminado. Ele também afirma que, mesmo que haja um pronome indefinido na frase como *alguém*, *ninguém*, ainda assim o sujeito é expresso, não se configurando caso de indeterminação. Assim, nos exemplos “*Alguém* me ensinará o caminho” e “*Ninguém* lhe telefonou”, os agentes, expressos pela voz ativa, são os pronomes indefinidos, em destaque. Para Cegalla (2008), a indeterminação do sujeito acontece em três modos. O primeiro é usando o verbo na *terceira pessoa do plural*, sem referência a qualquer agente já expresso nas orações anteriores, como em:

(17) “Na rua olhavam-no com admiração”

(18) “Bateram palmas no portãozinho da frente.”

O segundo é utilizando um verbo ativo na *terceira pessoa do singular*, acompanhado do pronome *se*, a exemplos de:

(19) “Aqui *se vive* bem.”

(20) “Devagar *se vai* ao longe”

(21) “*Trata-se* de fenômenos que nem a ciência sabe explicar.”

Já, o terceiro modo, quando se tem o verbo no infinitivo impessoal, exemplificados a seguir:

(22) “Era penoso *carregar* aqueles fardos enormes.”

(23) “É triste *assistir* a estas cenas repulsivas.

Rocha Lima (2011, p. 289) usa o critério de determinação para definir a indeterminação, assim, o sujeito pode ser determinado ou não. “ É determinado, se identificável na oração – explícita ou implicitamente; indeterminado, se não pudermos ou não quisermos especificá-lo”. Nas situações em que o sujeito é considerado indeterminado, a língua se vale de dois contextos<sup>38</sup>.

1. Empregar o verbo na 3ª pessoa do plural, sem referência anterior ao pronome *eles* ou *elas*, e a substantivo no plural;
2. Usá-lo na 3ª pessoa do singular acompanhado da partícula *se*, desde que o verbo seja intransitivo, ou traga complemento preposicional.

O primeiro contexto é exemplificado pelo gramático com “*Falam mal daquela moça.*”, “*Mataram um guarda.*”, enquanto, o segundo contexto é ilustrado pelas seguintes ações: *Vive-se bem aqui*” e “*Precisa-se de professores.*”

Conforme Faraco e Moura, (1997, p. 315) o sujeito indeterminado é aquele que “[...] não pode ser identificado nem pelo contexto nem pela terminação do verbo” e pode ocorrer nas situações oracionais em que há a presença de verbos na 3ª pessoa do plural, quando não é possível identificá-lo pelo contexto, a exemplo de:

(24) “Na sala *discutiam* agora a hora do enterro.” (Alcântara Machado)

A outra situação acontece com verbos na 3ª pessoa do singular, acompanhados da partícula *se*:

(25) “*Trata-se* de um médico com uma ótima imagem”

(26) “Não *se sabe* de um caso de assalto recente”.

---

<sup>38</sup> Grifos do autor

Há de se destacar que alguns gramáticos acabam por considerar o pronome substantivo indefinido como um elemento que indetermina o sujeito da oração:

(27) “*Tudo* assustava a pobre mulher”

(28) “*Ninguém* gostou daquele filme”.

Para o autor, uma análise apenas do ponto de vista sintático, consideraria as expressões *tudo* e *ninguém*, como casos de sujeito simples, mas, que do ponto de vista semântico, seriam indeterminados.

Faraco e Moura (1997, p. 316) afirmam que, quando se tem uma partícula *se*, pode se tornar mais difícil a classificação do sujeito e, eles comparam duas situações para ajudar na distinção, ilustradas por A e B:

A. *Discutiu-se o projeto. Discutiram-se os projetos.*

B. *Precisa-se de secretária. Precisa-se de secretárias.*

Com base nos exemplos dados em A e B, os autores afirmam que, no primeiro caso, a partícula *se* se comporta como uma partícula apassivadora e o verbo, que está na voz passiva sintética, concorda com o sujeito. Tal fato pode ser confirmado quando há a possibilidade de colocá-lo na voz passiva analítica, assim, as expressões *o projeto* e *os projetos* se tornam os sujeitos das frases. Tal situação não se aplica ao caso B, pois, embora o sujeito esteja também na voz ativa, a partícula *se* se comporta como índice de indeterminação do sujeito e, deste modo, o sujeito é classificado como indeterminado pois o sujeito permanece na terceira pessoa (P3), ou seja, “o verbo fica sempre na 3ª pessoa do singular” (Faraco; Moura, 1997, p. 317). Ainda sobre o caso A, os autores fazem uma crítica ao estilo coloquial do falante e até mesmo de registros literários encontrados em que a regra é infringida e citam “de longe se via os bichos bodejando no capim” de J. J. Veiga e “Aluga-se casas”, na Folha de S. Paulo.

Bechara (1966, p. 247) afirma que há três formas de classificar o sujeito como indeterminado:

1. verbo na 3ª pessoa do plural sem que haja referência, anterior ou seguinte, que sirva de sujeito:

(29) “Nunca me *disseram* isso. Onde *puseram* o livro?”

2. verbo no infinitivo ou na 3ª pessoa do singular com valor de 3ª pessoa do plural, sem referência que seja possível identificar um referente.

(30) “É bom resolver o problema.”

(31) “Diz que o fato não aconteceu assim. (diz = dizem)”

Para Bechara, esse segundo contexto, embora seja comum utilizar o verbo no infinitivo com a P3, uma quantidade maior de realizações acontece com verbos no plural. A terceira situação acontece quando há:

3. Verbo na 3ª pessoa do singular acompanhado do pronome *se*, que não seja seguido de um substantivo ou outro referente que sirva de sujeito do conteúdo predicativo. Nesse caso, o sujeito é indiferenciado e geral. Não somente, o “*se*” se comporta como um índice de indeterminação do sujeito ou pronome indeterminador do sujeito, como são os casos apresentados pelo autor:

(32) “Vive-se bem aqui.”

(33) “Lê-se pouco entre nós.”

(34) “Precisa-se de empregados.”

(35) “É-se feliz”.

#### 4.2.2 Outras concepções: determinação, indeterminação e indefinição para o estudo do sujeito

Nesta subseção são discutidos alguns conceitos do que pode ser considerado determinado ou não e outras definições que se julgam necessárias para o estudo da indeterminação do sujeito como é o caso da diferença entre indeterminação, indefinição. Para essa discussão, são utilizados, principalmente, os conceitos dados por Milanez (1982), já que essa autora serviu de referência para vários estudos sociolinguísticos sobre o fenômeno em questão.

Embora exista uma semelhança entre os termos indeterminação e indefinição, Milanez (1982) afirma que as duas podem confundir pois, ambas, expressam uma ideia não precisa. Diante disso, a autora expõe algumas diferenças para que a análise seja mais certa visto que as pessoas possam se sentir confusas para identificar os dois casos.

As características, segundo a autora, se baseiam, primeiramente por traços “mais humano” ou “menos humano”. Enquanto a indefinição possa ser considerada a partir de um elemento humano/não humano, a indeterminação só pode acontecer se apresentar traços [± humano], assim, só pode se referenciar aquilo que é humano. Um outro ponto que pode distinguir a indeterminação da indefinição são os recursos sintáticos. Sendo assim, a indeterminação se refere não só as formas verbais como também as lexicais, enquanto, a

indefinição se refere apenas às formas lexicais de terceira pessoa. Ao passo que a primeira tem caráter mais abrangente, a segunda, tem suas limitações.

Um outro ponto é a generalização. Enquanto esta é característica importante, a indefinição apenas apresenta uma possibilidade, assim, a indefinição pode indicar um referente individualizado ou generalizado, ainda assim, de maneira limitada. Para Milanez (1982, p. 39), “os recursos de indefinição sempre pressupõem um conjunto “fechado” de elementos que pode ser expresso ou na sua totalidade [...] ou no seu esvaziamento [...] ou parcialmente [...].

Outro aspecto que diferencia é que ela apresenta distinções morfológicas, a exemplo das marcas de singular e plural.

Milanez (1982, p. 80) afirma que:

grau mais alto de indeterminação seria expresso pelas três formas despronominalizadas ( $\emptyset+3^a$  p.s.), ( $\emptyset+3^a$  p.s.+se) + ( $\emptyset$ +infinitivo) as quais, conforme explicamos, distanciam-se de referências específicas a qualquer das três pessoas do discurso, uma vez que são marcadas quanto à pessoa.

Nesse contexto, as discussões sucedem-se não apenas do ponto de vista sintático, mas também do semântico e do discursivo, visto que é um fenômeno bastante complexo. Em sua *Gramática Pedagógica do Português Brasileiro*, Bagno (2011) concebe a indeterminação do sujeito como algo que não é estritamente sintático, mas, que também é semântico e afirma que:

Todos os estudos científicos empreendidos em torno desse tema, de Said Ali até hoje, têm se empenhado em demonstrar a urgente necessidade de interpretar o se nessas orações como um recurso de que a língua dispõe para indicar a *indeterminação do sujeito*. A nomenclatura que ainda usa conceitos como "se apassivador", "passiva sintética" e "passiva pronominal" é inteiramente descabida e tem de ser abandonada de uma vez por todas, junto com a concordância bizarra que ela implica. (Bagno, 2011, p. 807).

Bagno (2011) tece críticas às prescrições normativas e apresenta outras estratégias com a função de indeterminar o sujeito, mais inovadoras e que vão de encontro às formas cristalizadas nas gramáticas tradicionais, como é o caso do *verbo na não-pessoa do singular*, como em:

(36) “fale por si só”

O *verbo na não-pessoa do plural*, como em:

(37) “é um negócio sensacional, sabe **serviram** um chazinho especial também”

(NURC/POA/291);

*Eles*, em

(38) “[...] que para a mulher era muito difícil que **eles** quase não admitiam”  
(NURC/SP/360)

E, os *termos genéricos*, como em:

(39) “[...] então acha porque **a pessoa** trabalha com isso... é fácil de arranjar emprego [...]” (NURC/SP/360).

O autor ainda ressalta que:

A indeterminação do agente é um traço semântico, isto é, tem a ver com o signo linguístico e suas relações com o referente no mundo real. Quando não sabemos, não podemos ou não queremos enunciar esse agente, empregamos formas que expressam essa indeterminação — indeterminação que também empregamos para generalizar ações que, para nós, podem ser desempenhadas por qualquer pessoa. (Bagno, 2011, p. 803).

Ainda, sobre a forma que a Gramática Normativa discorre sobre os fenômenos da língua, Milanez afirma:

A GT não leva em conta a distinção entre os níveis sintático, semântico e pragmático da língua, daí o fato de determinados assuntos serem tratados ora num nível ora noutro, sem justificativas e de modo incoerente. É o que sucede, por exemplo, com o tradicional conceito de sujeito (sempre definido a nível semântico e interpretado a nível sintático, através da concordância), com o pronome se e com o próprio conceito de indeterminação. (Milanez, 1982, p. 21-22)

Para a alicerçar essas novas ideias, são tomadas, mais adiante, como referência também as gramáticas de Perini (2002), Neves (2011), Bagno (2011) e Castilho (2019), à medida que são discutidos alguns dos estudos sobre o fenômeno da indeterminação, (cf. seção 4.2.3) e, no capítulo 5 em que se discutem os aspectos metodológicos e, principalmente, quando as variantes são apresentadas.

#### 4.2.3 Estudos sobre a indeterminação do sujeito Brasil/Angola

Partindo do pressuposto de que toda língua apresenta variações que não são aleatórias e sim sistemáticas, em diferentes comunidades de fala, é necessário analisar os contextos de algumas escolhas feitas pelos falantes por uma determinada variante em detrimento de outras que vão muito além daquelas formas cristalizadas nas gramáticas tradicionais. Com o fenômeno

variável da indeterminação não é diferente, portanto, é importante fazer um levantamento de alguns estudos já realizados em território brasileiro e angolano para que seja possível observar como o fenômeno se comporta nas distintas comunidades, já que as formas mais utilizadas por elas nem sempre são as mesmas e nem suas frequências.

#### 4.2.3.1 Estudos Linguísticas em amostras de fala no Brasil

Milanez (1982), em seu trabalho “Recursos de indeterminação do sujeito”, utilizou o *corpus* de fala de 20 informantes de São Paulo, do Projeto NURC/SP e, ao utilizar a perspectiva gerativista transformacional, buscou descrever as estratégias para indeterminar os sujeitos e seus resultados extrapolaram a quantidade de formas apresentadas pelas GTs: *a gente, você, eles, eu, nós, o/um indivíduo, o/um sujeito, o/uma pessoa, o/um cara, Ø + V3PS, Ø + 3VPP, Ø + verbo + se* e o *Ø + infinitivo*. Além das formas nominais, a autora subdividiu as estratégias de indeterminação, em dois outros grupos: aquelas que tinham indeterminação com sujeito lexical (*a gente, você* e *eles*), e aquelas sem sujeito lexical (*Ø + 3VPS, Ø + 3VPP, Ø + verbo + se*, e *Ø + infinitivo*). O grupo com maior porcentagem de ocorrência foi a com sujeito lexical preenchido (62%), com o recurso *a gente* (Oc. 361), em destaque, e, o restante foi das formas não pronominalizadas. A quantidade de ocorrências do *a gente*, foi seguida pela partícula *se* (Oc. 205) e *você* (Oc. 168).

Rolleberg (1991) desenvolveu sua pesquisa *Os pronomes pessoais e a indeterminação do sujeito na norma culta de Salvador*, em cooperação com alunos bolsistas de IC, utilizando os dados de fala de pessoas que tinham o ensino superior completo, e que foram coletados, ainda na década de 1970, pelo Projeto NURC/SSA. Ao considerar as formas pronominais como recursos de indeterminação do sujeito, seu trabalho teve como objetivo inicial compreender como a indeterminação do sujeito estava presente nas gramáticas para então observar como os pronomes pessoais estavam sendo utilizados em Salvador como um recurso de indeterminação.

A autora parte do pressuposto de que o estudo do sujeito indeterminado deve se pautar em critérios não só semântico-pragmáticos, mas também, sintáticos. Ela relacionou suas estratégias de indeterminação com outras duas variáveis, a *categoria do texto* (EF<sup>39</sup>/DID<sup>40</sup>) e a *faixa etária dos informantes* (Faixas 1 e 3, conforme estabelecido pelo NURC). Para a realização de sua análise, Rolleberg (1991), se baseou nas categorias de pessoas de

<sup>39</sup> EF- Elocução Formal;

<sup>40</sup> DID- Diálogo entre Informante e Documentador

Benveniste (1976) – teoria da enunciação - e, a propriedade semântica da referencialidade, para as questões de indeterminação.

Não houve registro das formas verbais despronominalizadas (*terceira pessoa do singular acompanhada do se, terceira pessoa do plural e verbo infinito impessoal*), enquanto, no grupo dos pronominais foram registradas as formas *você* com 35,54%, *nós* com 32,23%, *a gente* 25,12% e *eles*, em menor quantidade, com 7,11% dos dados encontrados. Seus resultados mostraram que a forma *você* foi a mais utilizada pelos falantes mais jovens (Faixa 1) e cultos de Salvador, com prevalência na categoria de texto *DID*, e a forma *eles* foi a mais utilizada pelos falantes mais velhos (Faixa 3).

Menon (1994), realizou uma pesquisa sociovariacionista em sua tese intitulada “Analyse sociolinguistique de l’indétermination du sujet dans le portugais parlé au Brésil, a partir des données du NURC/SP” utilizando um *corpus* de fala urbana de São Paulo com 68 entrevistas. A indeterminação do sujeito para Menon (1994) está sempre ligada à indeterminação do referente e se expressam de uma forma mais ampla daqueles previstos pela GT. A autora considerou as seguintes formas: *a gente, ele(s), eu, formas nominais, nós, se, você(s), voz passiva sem agente, voz passiva sintética, verbo na 3ª pessoa do singular e verbo na 3ª pessoa do plural*. Esse grande número de variantes para indeterminar o sujeito diferiu dos estudos anteriores realizados por Omena (1987), que focou na comparação do uso de *nós* e *a gente* e das variantes *você, nós* e *a gente* como indeterminação e de Cunha (1993), que se restringiu ao estudo do caráter indeterminador dos pronomes *nós, você* e *a gente* e Milanez (1982) que acrescentou mais variantes que suas antecessoras como: *a gente, você, eles, ØV3PS, Ø+INF, ØV3PP, se, eu* e alguns *sintagmas nominais* como: *a pessoa, o indivíduo, o sujeito e o cara*.

Os dados coletados por Menon foram analisados utilizando as seguintes variáveis extralinguísticas: gênero/sexo (M/F), tipos de entrevistas (EF, DID e D2<sup>41</sup>) e faixa etária (25-35; 36-55 e acima de 56 anos). Alguns dos resultados obtidos pela autora foi a comprovação de sua hipótese de que os falantes usam mais recursos para indeterminar o sujeito do que aquelas presentes nas GTs e que são ensinadas nas escolas. Em relação às variantes linguísticas, os resultados da autora apresentaram: (i) o *a gente*, devido a um processo de gramaticalização, tornou-se um pronome genérico, depois indeterminado e, ainda utilizados como valor de 1ª pessoa, principalmente plural; (ii) o termo *a pessoa* passou a ser um substitutivo do *a gente* pelos falantes como se estivesse ocupando um espaço vazio; (iii) o pronome *se* se caracteriza

---

<sup>41</sup> D2- Diálogo entre dois informantes

como um sujeito de pleno direito;(iv) *você* em concorrência com a forma *tu* tornou-se um pronome pessoal de segunda pessoa, além de portar a função de indeterminação do sujeito.

Ainda como resultados encontrados por Menon (1994), foi a preferência das mulheres para as formas inovadoras, enquanto os homens utilizaram as estratégias mais formais de indeterminação. Destaca-se o papel dos mais jovens ao preferirem as formas *eu*, *eles* e *você* enquanto os falantes mais velhos utilizavam mais o *verbo na terceira pessoa do plural* e a *voz passiva sintética*, enquanto o *verbo na terceira pessoa do singular* foi utilizado pelas três faixas etárias analisadas, indicando estabilidade na comunidade investigada.

O estudo de Setti (1997) buscou investigar o comportamento linguístico dos falantes em relação à indeterminação do sujeito nas três capitais da região Sul (Curitiba, Porto Alegre e Florianópolis) e, assim como Menon (1994), ela fez algumas substituições de variantes em enunciados da amostra e comprovou a intercambialidade das variantes mantendo o mesmo valor da indeterminação, pois “percebemos que os falantes utilizam variadas formas para representar o sujeito, numa mesma oração, e o caráter genérico do referente permanece.” (Setti, 1997, p. 18). São exemplos da sua amostra:

#### **VOCÊ x O CARA**

"Então a gente vê, se **você** não for comunicativo, **você** não vai ter nada. " (CBA/01/1074)

Então a gente vê, se **o cara** não for comunicativo, **o cara** não vai ter nada.

#### **ØV3PS x A GENTE**

"Depois **Ø** tinha que vim a pé de lá pra trazer pra cá uns ("trequinhos") de carroça, né?" (CBA/01/0167)

Depois **a gente** tinha que vim a pé de lá pra trazer pra cá uns ("trequinhos") de carroça, né? (Setti, 1997, p. 19-20)

A essa intercambialidade das variantes (pares mínimos<sup>42</sup>), em que usar uma variante no lugar da outra sem que o sentido fosse alterado, configura uma forma que justifica as treze variáveis elencadas para o estudo. A autora considerou como sujeito indeterminado o contexto de não existir possibilidade ou quando não há interesse em identificar o referente. Setti (1997) utilizou os mesmos condicionamentos linguísticos de Menon (1994) e as estratégias selecionadas por ela para a indeterminação do sujeito foram: a gente, eles, eu, formas nominais,

---

<sup>42</sup> Os pares mínimos “são tipos especiais de variantes passíveis de serem usadas exatamente num mesmo contexto, uma no lugar da outra”. (Setti, 1997, p. 20)

nós, se, tu, você(s), voz passiva sem agente, voz passiva sintética, verbo na 3ª pessoa do singular e verbo na 3ª pessoa do plural e baseadas no estudo prévio de Menon (1994). Diferentemente de Menon (1994), Setti (1997), analisou a variante *tu* pois os dados do NURC- SP não tiveram qualquer ocorrência dessa variante com indeterminação do sujeito. A amostra de Setti (1997) foi constituída de um total de 72 entrevistas sendo que 24 entrevistas para cada uma das capitais analisadas, 08 informantes por nível de escolaridade (total de 24), 12 por sexo (total de 36) e 12 por faixa etária (total de 36).

Para a realização desse estudo, os dados do Projeto VARSUL<sup>43</sup> foram utilizados e as análises se pautaram nos condicionamentos linguísticos e sociais, em que algumas das formas elencadas por Setti foram submetidas às rodadas estatísticas como variáveis dependentes. Segundo a autora, “... foi sob o filtro desses condicionamentos que se pôde observar a variação relacionada à indeterminação do sujeito. A intensidade de uso de cada variante foi resultante da influência dos fatores extralinguísticos envolvidos.” (Setti, 1997, p. 107).

Não diferente de outras pesquisas, Setti (1997), comprova que há mais possibilidades para se indeterminar o sujeito do que as prescritas nas GTs, e afirma que “apesar de diferenças geográficas, etárias, culturais e outras englobadas em cada estudo, os falantes têm em comum o contato com um universo de recursos linguísticos amplo” (Setti, 1997, p.108). Em seu estudo, a hipótese levantada para os tempos e modos verbais como favorecedores para os recursos de indeterminação são também comprovadas, principalmente, quando se trata do presente do indicativo. Sobre a variável escolaridade, os comportamentos em relação as homens e mulheres foram parecidos e mostraram um certa neutralidade. Os falantes menos escolarizados preenchem mais o sujeito do que os mais escolarizados (tendência a utilizar o  $\emptyset+V3PP$ ). Em relação à variável sexo/gênero, as mulheres se mostraram mais conservadoras ao passo que os homens utilizaram as formas mais inovadoras.

No que concerne às localidades, os dados mais gerais para as capitais foram as formas  $\emptyset+V3PS$ , *a gente* e *você* com maiores frequências relativas e, em Porto Alegre e Florianópolis houve a predominância da forma *tu*, exclusivamente, conforme a autora. Já as formas menos frequentes foram *vocês* e *VPASSINT*. Essa última recebe uma ressalva da autora ao dizer que parece estar em desuso na língua oral, diferentemente do que é visto nas gramáticas, e que

---

<sup>43</sup> O projeto VARSUL (Variação Linguística na Região Sul do Brasil) estuda o português falado e escrito de áreas do Sul do Brasil, com a parceria da Universidade Federal do Rio Grande do Sul, da Pontifícia Universidade Católica do Rio Grande do Sul, da Universidade Federal de Santa Catarina e da Universidade Federal do Paraná. O VARSUL estratifica a amostra da seguinte maneira: o sexo (M/F), a idade (A: 25 a 50 anos; B: mais de 50 anos) e o grau de instrução (PRI: primário/course elementar; GIN: ginásio/5ª a 8ª série; SEC: secundário 2º grau).

reflete no ensino, agravando o distanciamento do que de fato é ensinado com o uso real pelos falantes.

Em sua dissertação de Mestrado, Santana (2006) analisou as estratégias de indeterminação do sujeito no discurso oral de duas comunidades rurais afrodescendentes do semiárido baiano: Piabas e Bananal/Barra dos Negros, na região de Rio de Contas. A amostra foi constituída de 24 entrevistas, sendo 12 de cada comunidade e estratificada em Faixa 1, 2 e 3 entre homens e mulheres. Os *corpora* utilizados fazem parte do projeto “A Língua Portuguesa no Semiárido Baiano” desenvolvido por pesquisadoras da UEFS/BA, entre os períodos de 1994 e 2000. As possibilidades de indeterminação do sujeito elencadas foram: *a gente*, *nós*, *você*, *eles*,  $\emptyset+V3PP$ ,  $\emptyset+V3PS$ ,  $\emptyset+VINF$  e  $\emptyset+V+SE$ . Em relação aos fatores linguísticos, foram utilizados o tempo e modo verbal, tipo de verbo, tipo de oração, preenchimento do sujeito, grau de indeterminação, forma antecedente e mudança/manutenção do referente, e os extralinguísticos foram o gênero, faixa etária e localidade.

Os resultados encontrados por Santana (2006) mostraram que de fato existem muito mais formas utilizadas pelas comunidades do que aquelas previstas nas gramáticas e que essas últimas não foram as mais utilizadas. Os recursos mais utilizados pelas comunidades foram as formas *a gente* com 45% dos dados e  $\emptyset+V3PS$  com 20%. As seguintes variáveis dependentes apresentaram os seguintes valores: *eles* (10,5%), *nós* (9%),  $\emptyset+V3PP$  (6%), *você* (5%),  $\emptyset+VINF$  (4%) e  $\emptyset+V+SE$  (0,5%). Deste modo, em seus dados, 30% dos recursos de indeterminação na amostra foram oriundos de formas sem sujeito lexical e a maioria, 70% dos pronomes pessoais, em especial, da forma *a gente*. Em relação ao *nós* X *a gente*, há uma grande retração no uso do *nós* que pode estar relacionada à redução nos paradigmas flexionais do PB como atestam Monteiro (1994), Duarte (1995) e Menon (1994) conforme explica a autora.

Tanto os fatores linguísticos e extralinguísticos exerceram influências na escolha da forma pelos falantes e, foram feitas análises binárias confrontando as variáveis dependentes. Conforme a autora, alguns grupos de fatores eram frequentemente selecionados para todas as variantes como a forma antecedente, com exceção da variante  $\emptyset+V3PP$  e corrobora os resultados de Cunha (1993), não somente, quando esse fator foi cruzado com a mudança/manutenção do referente, os resultados indicaram que a mesma forma é sempre usada quando a referência é a mesma. Em relação ao grau de indeterminação, *nós* e *a gente* foram mais usadas para indicar indeterminação parcial com referência implícita ou explícita no contexto, enquanto *eles* foi frequente quando houve indeterminação parcial com referência explícita no contexto; houve uma indeterminação completa (mais genérica, referindo-se a qualquer pessoa do discurso) nos contextos de *você*,  $\emptyset+V3PS$ ,  $\emptyset+V3PP$  e  $\emptyset+VINF$ .

Sobre o tempo e modo verbal, com todas as variantes, o mais utilizado pelos falantes é o presente do indicativo. Os dados de Santana (2006) se assemelham ao encontrado por Setti (1997) e Godoy (1999). Sobre os fatores extralinguísticos, homens e mulheres se comportaram de maneiras distintas. As mulheres usaram mais a forma *a gente*, enquanto os homens variavam mais os recursos de indeterminação utilizando *você*,  $\emptyset+V3PS$ ,  $\emptyset+VINF$ . Quanto à faixa etária, as faixas 2 e 3 favoreceram mais o uso das formas indeterminadas, com prevalência das formas *a gente* e *você* pelos informantes da faixa 2 e  $\emptyset+V3PS$  pelos informantes da faixa 3. Em relação à última variável extralinguística, as localidades investigadas apresentaram resultados diferentes. A comunidade de Piabas favoreceu as formas mais conservadoras (*nós* e  $\emptyset+V3PP$ ), enquanto Bananal/Barra dos Negros utilizou a forma mais inovadora *a gente*. Sobre essa diferenciação, a autora explica o fato de a comunidade Piabas ser mais isolada pois não é passagem nem entroncamento rodoviário. As pessoas da comunidade até vão à cidade, mas devido à falta de atrativos no local, quase não recebem visitas.

Os resultados de Santana (2006) mostraram que, em relação as formas de indeterminação do sujeito, há uma distância muito grande do uso tradicional das gramáticas e o que de fato ocorre na língua falada nas comunidades rurais. É comprovado que esse fenômeno está intimamente ligado com os níveis semânticos e discursivos já que muito se deve ao contexto.

Ponte (2008), analisa as estratégias de indeterminação do sujeito do português popular o interior do estado da Bahia com amostras de pesquisas realizadas nas comunidades de Barra e Bananal (Rio de Contas), Sapé (Valença), Helvécia (Nova Viçosa) e Cinzento (Planalto) e nas zonas rural e urbana de Santo Antônio de Jesus. Foi utilizado o *Corpus* Base do Português Afro-Brasileiro do Estado da Bahia além de 24 entrevistas coletadas em Santo Antônio de Jesus do Acervo de Fala Vernácula do Português Afro-Brasileiro do Estado da Bahia, sendo que dois informantes das três faixas etárias foram selecionados, um homem e uma mulher de cada uma das quatro comunidades pertencentes ao Acervo. As entrevistas foram realizadas entre 1992 e 2002 é feito um estudo em tempo aparente. Os informantes foram distribuídos por localidade, gênero e faixa etária, escolaridade e estada fora da comunidade por seis meses ou mais. No que concerne as variáveis linguísticas foram analisados: (i) realização fonética ou não do sujeito indeterminado; (ii) desinência de pessoa e número presente ou não na forma verbal em foco; (iii) tipo de verbo; (iv) tipo de frase; (v) nível de referencialidade do agente; (vi) modo verbal; (vii) inclusão do falante no universo de referência do sujeito indeterminado; (viii) paralelismo formal.

A autora utiliza as hipóteses da TLI e da polarização sociolinguística para explicar os aspectos morfossintáticos que se distanciam ou aproximam do uso da língua culta nessas

comunidades interioranas. Nessa pesquisa baseada na teoria e metodologia da Sociolinguística Variacionista, a autora levanta algumas hipóteses para o português popular no interior do estado da Bahia em relação à indeterminação do sujeito: a) alta realização dos sujeitos pronominais devido à redução do paradigma flexional dos verbos que resultara da aquisição defectiva do português; baixa ocorrência do *se*, principalmente nas comunidades afro-brasileiras devido à redução da morfologia da língua; alta frequência das formas menos marcadas, como é o caso da categoria vazia ( $\emptyset$  + V3PS) e a existência de um *continuum* linguístico oriundos dos processos de TLI.

Conforme já foi investigado por vários autores, foi verificado que nessas comunidades há também uma variedade de estratégias se comparadas àquelas presentes nas GTs. Foram estudadas as seguintes variantes: os pronomes indeterminadores nós, a gente, você e eles e as formas verbais  $\emptyset$  + V3PP e  $\emptyset$  + V3PS. A análise contou com 3.830 ocorrências sendo o pronome *a gente* obteve 47%, seguido da forma  $\emptyset$  + V3PS, com 25%, os mais produtivos. Você resultou em 9% e as demais, valores abaixo de 6%.

Sobre os quatro níveis de referencialidade, o nível genérico, prevaleceram as formas: *a gente*, *você* e  $\emptyset$  + V3PS. Já os níveis grupo parcialmente definido e eu ampliado, destacam-se os pronomes P4 (*a gente* e *nós*) e a forma  $\emptyset$  + V3PS. Quanto ao nível do agente específico indefinido, o predominou P6 (*eles* e  $\emptyset$  + V3PP). De todas as formas destacadas, a que teve maior relevância foi a estratégia  $\emptyset$  + V3PS, presente nos 4 níveis de referencialidade analisados para o sujeito indeterminado. Em relação aos contextos linguísticos, a forma antecedente foi a única selecionada pelas 6 estratégias de indeterminação indicando que uma forma utilizada tende a se repetir na cadeia discursiva. O modo verbal foi selecionado pelas variantes *nós* e  $\emptyset$  + V3PP, indicados por um fato e os irrealis pelo *você* e *eles*. O caso de *você* é relevante na amostra, visto que os dados apontaram que esse pronome quase nunca excluiu o falante.

Considerando os contextos sociais, o mais relevante foi a localidade do informante em relação as 6 variantes atestadas mostrando um *continuum* linguístico da forma  $\emptyset$  + V3PS, pois apresenta maior funcionalidade nas comunidades afro-brasileiras do que nas outras localidades da amostra. Sobre a variável estada fora da comunidade, as formas *a gente*,  $\emptyset$  + V3PP e  $\emptyset$  + V3PS foram as escolhas dos falantes. Ponte (2008) afirma que os seus resultados corroboram com a hipótese da difusão do português padrão partindo dos grandes centros urbanos já que *a gente* e a estratégia padrão  $\emptyset$  + V3PP são favorecidos pelos falantes que saíram da comunidade enquanto a forma  $\emptyset$  + V3PS, não padrão, é mais utilizada por aqueles que não tiveram contato com outras localidades. Os falantes mais escolarizados, bem como as mulheres favoreceram a estratégia do pronome *nós*. Nas comunidades interioranas pesquisadas por Ponte (2008), a

variante padrão, Ø + V3PS + SE, quase não é usada pelos informantes. Na pesquisa, a variável faixa etária foi relevante considerando a substituição do pronome *nós* pelo *a gente*.

Ponte (2008) conclui que existem fortes marcas da aquisição imperfeita no PVB e o afrouxamento dos padrões linguísticos pelas camadas médias e altas, devido ao nivelamento linguístico que chegou as variedades populares do PB. As razões são as mais variadas e já discutidas na Seção 1 desta tese, a exemplo da recente e intensa urbanização, da escolarização e dos meios de comunicação de massa. Para a autora, as características linguísticas consideradas mais crioulizantes podem só ser encontradas em poucas comunidades rurais afro-brasileiras isoladas.

Carvalho (2010) em sua dissertação de mestrado intitulada *Você, a gente et alia indeterminam o sujeito em Salvador*, investigou as estratégias de indeterminação do sujeito utilizando 44 inquéritos, 32 pertencentes ao Programa de Estudos sobre o Português Popular Falado de Salvador (PEPP) e 12 do Projeto NURC/Salvador. O estudo se caracterizou como um estudo sincrônico e se pautou nos pressupostos teórico-metodológicos laboviano e no funcionalismo (gramaticalização). Carvalho (2010) controlou as seguintes variáveis linguísticas: *forma antecedente*, *mudança/manutenção do referente*, *tempo e modo verbais*, *tipos de oração*, *tipo de verbo*, *preenchimento do sujeito* e *grau de indeterminação*. As variáveis extralinguísticas se dividiram em *sexo/gênero*, *escolaridade* e *faixa etária*.

Ao partir da hipótese de que os falantes usam muito mais estratégias para indeterminar o sujeito que não sejam as clássicas presentes nos manuais tradicionais, ele controlou as seguintes variantes: *a gente*, *eles*, *eu*, formas nominais, *nós*, *você*, [Ø+V+SE], [Ø+V3PP], [Ø+V3PS], [Ø+VINFIN], [VPSA] e [VPASSINT]. Assim, as 12 variantes elencadas foram divididas em 2 grupos, com sujeito lexical preenchido e sem sujeito lexical, já proposto por Milanez (1982).

Carvalho (2010) parte do pressuposto de que a indeterminação do sujeito é uma indeterminação do seu referente (agente da ação), mas que se encontra em um contexto extralinguístico e corrobora o que Menon (1994, 2006) já tinha conceituado de referente extralinguístico. Essa posição dos autores reforça a ideia que vai além da simples definição de sujeito indeterminado ofertado pelos compêndios gramaticais, mas que ao mesmo tempo, também é possível analisar todo o contexto da interação verbal presente no ato comunicativo.

Os resultados de Carvalho ao analisar a fala de pessoas em Salvador (2010) mostram as formas *você*, *a gente* e as *formas nominais* como as mais presentes nos *corpora* investigados, estratégias bem diferentes das encontradas nas GTs. Considerando as variáveis linguísticas, as formas pronominais, a exemplo de *você*, *a gente*, *nós*, *eles* e *eu* prevaleceram das não-pronominais e foi motivada, principalmente, pela variável *forma antecedente*, sobretudo

quando houve a manutenção do referente, quando cruzado com a variável *mudança/manutenção do referente*. Interessantemente, a variável *forma antecedente* ainda foi selecionada nas rodadas que envolveram as estratégias não-pronominais como também em muitas outras rodadas realizadas separadamente pelo autor e, quase sempre, atrelada à variável *mudança/manutenção do referente*.

Já, em relação às variáveis extralinguísticas, o nível de escolaridade se mostrou produtivo com destaque para o Ensino Superior e sobre isso, o autor destaca:

“O Ensino Superior contribui para que haja o fortalecimento no uso dessas estratégias, ou seja, quanto mais escolaridade tem o falante, maior é o seu leque de opções linguísticas para marcar a indeterminação do sujeito em Salvador.” (Carvalho, 2010, p. 189-190)

Já, a pesquisa sociolinguística feita por Assunção (2012) contou com dados de entrevistas realizadas do tipo DID entre (2009 e 2012) do projeto “A língua portuguesa falada no semiárido baiano” da UEFS ao analisar a fala de 24 feirenses de ambos os sexos, com escolaridades fundamental e superior de três diferentes faixas etárias. O objetivo era identificar na variedade feirense formas variadas para indeterminar o sujeito através de fatores linguísticos e extralinguísticos. Segundo os resultados da autora, os falantes feirenses utilizam as formas pronominais *eu, você, a gente, nós, eles*, formas nominais,  $\emptyset+3PS$  e o  $\emptyset+VIN$  para indeterminar o sujeito, além das formas legitimadas pelas GTs com preferências pelas variantes *a gente e você*. Não foi encontrada a presença da variante *tu* no *corpus* feirense com referência indeterminadora. Nos resultados de Assunção (2012) as formas consideradas padrão pelas GTs,  $\emptyset + VERBO+SE$  e  $\emptyset V3PP$  não foram significativas na amostra. Essa última,  $\emptyset V3PP$ , por ser considerada mais popular, os poucos dados encontrados foram realizados pelos falantes menos escolarizados.

Ao considerar a regra de aplicação *você*, as variáveis linguísticas mais significativas observadas por ela foram o tipo de oração, tipo de verbo, preenchimento do sujeito. O tempo e o modo verbais também foram relevantes – futuro do subjuntivo com P.R. 0.82 e infinitivo, P.R. 0.76 e presente do indicativo (0.56). Em relação aos fatores extralinguísticos, todos os grupos foram selecionados na seguinte ordem – faixa etária (os mais jovens, P.R. 0.66), sexo/gênero (homens, P.R. 0.70) e escolaridade (ensino superior, P.R. 0.62). Nessa pesquisa, a autora considerou *você* como a variante mais inovadora utilizada pelos jovens e seus dados mostraram uma mudança em curso. Verificou-se também que, sobre a variável escolaridade, embora a indeterminação do sujeito esteja presente nos currículos escolares, ao mesmo tempo,

não sofre a atenção disciplinadora da escola, já que foi usada muitos falantes do ensino superior. Assunção (2012) conclui afirmando que de fato há mais possibilidades de se indeterminar o sujeito além daquelas previstas nas GTs, e estão condicionadas mais aos fatores extralinguísticos.

Souza (2014) também realizou um estudo variacionista na Bahia analisando as formas de indeterminação do sujeito na fala de moradores da zona rural da Matinha (Feira de Santana) e de Casinhas (Jeremoabo). A pesquisa utilizou 24 inquéritos do tipo DID, 12 para cada localidade, gravados e transcritos pelo projeto “A língua portuguesa no semiárido baiano”, com sede no Núcleo de Estudos de Língua Portuguesa (NELP) da UEFS. Souza (2014) considerou quatro variáveis linguísticas (tipo de verbo, tempo e modo verbal, paralelismo e tipos de oração) e quatro variáveis sociais (sexo/gênero, faixa etária, escolaridade e localidade). Os 2745 dados encontrados foram distribuídos conforme as diferentes variantes. As variantes controladas pela pesquisadora se assemelharam a de outras pesquisas como *a gente, nós, você, eles, Ø+V3PP, Ø+V3PS, Ø+V+SE*, como também das formas nominais.

A autora não considerou a variável *grau de indeterminação* como descrito por Milanez (1982) e utilizado nas pesquisas de Carvalho (2010), Santana (2006), Assunção (2012), Dias (2017), entre outros, assumindo a posição de que quando se trata de estratégias de indeterminação, não é possível observar maior ou menor grau de indeterminação, já que não é possível identificar nenhuma referência de sujeito, mesmo que alguns recursos apresentem marca morfológica de pessoa. Como também não destacou a *mudança/manutenção do referente* já que argumenta que se há indeterminação do sujeito, o referente é inexistente, portanto, não há como controlar se mudou ou não algo se ele não existe.

Os resultados obtidos por Souza (2014) indicaram que nas duas comunidades os informantes também usaram muito mais estratégias daquelas previstas pelas GTs e, que o pronome *a gente* (41%) e o *Ø+V3PS* (28%) foram os mais utilizados pelos falantes cujos resultados também se assemelham aos de Santana (2006) e Assunção (2012). Para as variantes mais tradicionais, o *Ø+V+SE* e *Ø+V3PP* as ocorrências foram ínfimas e colaboram com os resultados da pesquisa realizada por Santana (2006) nas comunidades rurais do semiárido baiano, Piabas e Bananal. A variável *formas nominais* (17%) foi a terceira variante mais usada e as formas encontradas na amostra foram *o cara, o camarada, o homem, a pessoa, o sujeito, a turma, o pessoal, nego, o cabra* (esse presente apenas em Casinhas).

Souza fez várias rodadas com as diferentes variantes, sendo que as variáveis linguísticas mais significativas foram o tempo/modo verbal e o paralelismo. Não somente, a aplicação da regra *a gente* foi favorecida diante de verbos no presente do indicativo, no futuro do presente

do indicativo e após forma idêntica (paralelismo). Houve também uma maior aplicação da regra com verbos intransitivos e, em orações subordinadas. Em relação aos fatores sociais, o grau de escolaridade, o sexo e a localidade dos informantes se sobressaíram, desse modo, o *a gente* foi mais utilizado por mulheres, da faixa 3 e que possuíam 2 a 3 anos de escolarização. O comportamento da variante “a gente” mostra um processo de mudança em curso visto que, cada vez mais, os falantes a utilizaram como uma forma indeterminadora. Fica evidente na pesquisa de Souza (2014) que existem muitas outras possibilidades de indeterminação que não são apresentadas pelas gramáticas, por isso, a autora mostra a necessidade de reformulação nos manuais de gramática.

Mendonça e Nascimento (2015) analisam o fenômeno da indeterminação do sujeito ao analisar a amostra Rede Social de Informantes Universitários de Itabaiana/SE (Araújo; Santos; Freitag, 2014) com o objetivo de verificar a correlação entre as estratégias de indeterminação do sujeito e o sexo/gênero. Deste modo, as autoras buscaram identificar as especificidades de cada recurso, não somente através do controle das relações de sexo/gênero, simétricas ou assimétricas, mas também fator pragmático distância social (expressão de polidez). Mendonça e Nascimento (2015) justificam o seu posicionamento para a realização desse estudo com base em Brown; Levinson, 2011[1987] ao considerar a indeterminação do sujeito uma estratégia de polidez, visto que acreditam que os recursos para indeterminar usados pelas pessoas podem interferir, impedindo, atenuando ou reparando possíveis ameaças à face locutor ou do interlocutor.

Ao observar o fenômeno do ponto de vista semântico-pragmático, as autoras afirmam que a indeterminação ocorre nas situações comunicativas e em contextos específicos, por isso, analisar o contexto em que ela se encontra é fundamental. As variáveis analisadas pelas autoras foram divididas em três categorias: a dos sintáticos, as previstas pela gramática normativa - *ØV3P, construções com se, infinitivo* -, a dos pronominais - *você, a gente, eles, nós, eu* - e, por fim, as nominais - *a pessoa e o cara*. Os dados apontam as estratégias nominais como as maiores favorecedoras correspondendo a 71% dos dados levantados, seguidos pelos sintáticos e pelas formas nominais com 25% e 4%, respectivamente. As autoras afirmam a importância de analisar o sexo/gênero como uma variável que ajuda a identificação de recursos utilizados já que podem ser característicos de uma preferência/escolha por algumas das formas elencadas, como o sexo feminino (*a gente*), e se baseiam nos estudos prévios de Setti (1997), Godoy (1999), Assunção (2012), Santana (2014), Souza e Oliveira (2014), entre outros para a realização de seu trabalho.

Guimarães (2018), em sua pesquisa geosociolinguística, objetivou identificar as formas de indeterminação mais recorrentes em 7 cidades<sup>44</sup> baianas das sete mesorregiões do estado, comparar as estratégias de indeterminação encontradas e identificar quais variáveis linguísticas e extralinguísticas favoreciam o uso de determinadas variantes. O corpus utilizado foi composto de 28 entrevistas do projeto Atlas Linguístico do Brasil<sup>45</sup> (ALiB) distribuídos em localidade, sexo e faixa etária. Em relação à escolaridade, só foram consideradas o nível fundamental, pois não há entrevista do Ensino Superior fora das capitais brasileiras pelo projeto supracitado. A análise contou com 811 dados após a entrevista em formato perguntas pré-elaboradas baseadas no questionário fonético-fonológico e semântico-lexical proposto pelo ALiB.

As variantes dependentes foram alocadas em três grupos: o primeiro, *gramaticais e suas variações*: (i) verbo na 3ª pessoa de singular, acompanhada do ‘se’, (ii) verbo no infinitivo, (iii) o verbo na 3ª pessoa do singular sem o ‘se’, (iv) 3ª pessoa do plural; o segundo, formas nominais, se subdividiram em 3 grupos menores: formas nominais A (artigo acompanhado de substantivo ou pronome indefinido), formas nominais B (pronome indefinido seguido de substantivo ou de pronome interrogativo), formas nominais C (estruturas formadas por um único elemento - um substantivo, um artigo, um pronome demonstrativo ou indefinido); o terceiro, *formas pronominais*: nós, a gente, você, eles.

As variáveis linguísticas controladas foram: a presença ou não de um sujeito lexical, o grau de indeterminação, tipo de verbo, tempo/modo e forma verbal, tipos de frase e o paralelismo sintático. Os resultados mais gerais apontam que há mais formas de indeterminação do que aquelas prescritas pela tradição gramatical, por exemplo o verbo na 3ª pessoa do singular desacompanhado da partícula ‘se’ e verbos no infinitivo. Os dados do Grupo 1 foram os que apresentaram um maior percentual (49,7%), seguidas das nominais e pronominais com 32,7% e 17,6%, respectivamente. No grupo 1, em relação as duas variantes mais encontradas, apenas a faixa etária foi selecionada com maior favorecimento dos falantes da faixa etária 2 (verbo na 3ª pessoa do singular sem ‘se’) e uma tendência ao desfavorecimento entre falantes da faixa etária 1 (preferência por verbos na 3ª pessoa do plural).

A autora afirma que, embora a variável ‘Cidades/mesorregiões’ não tenha sido selecionada, foi possível perceber uma tendência para um maior uso da forma 3ª pessoa do singular com ‘se’ com valores acima de 65% em todas as cidades analisadas. No grupo 1,

---

<sup>44</sup> As cidades (mesorregiões) investigadas foram: Barreiras (Extremo Oeste), Barra (Vale do São-Francisco), Vitória da Conquista (Centro-Sul), Ilhéus (Sul), Irecê (Centro-Norte), Alagoinhas (Nordeste) e Salvador (Região Metropolitana)

<sup>45</sup> A metodologia do ALiB compreende: Faixa 1 (18-30 anos), Faixa 2 (50-65 anos); Sexo: Homem/Mulher; Escolaridade: Nível fundamental/Nível superior.

nenhuma das variáveis linguísticas foram selecionadas. Sobre as variantes do Grupo 2, o subgrupo mais favorecedor foi das variantes cuja estrutura morfossintática era de um artigo acompanhado de substantivo ou pronome indefinido (55,8%), a exemplo de *as pessoas, os outros, o povo, o pessoal, a maioria, o(s) cara(s) e a turma*. Quando as cidades/mesorregiões, três se distinguiram das demais (Vitória da Conquista, Barreiras e Barra – parte oeste do estado), favorecendo as formas nominais do grupo A e as demais desfavorecendo.

A estratégia pronominal *a gente* chegou a quase 63% da amostra e os resultados de Guimarães (2018), resultados similares de Ponte (2008) e Souza (2014), ultrapassando a forma *nós*, com exceção de Barreiras que não prevaleceu o *a gente* tendo 36,4% dos dados, apenas.

#### 4.2.3.2 Estudos Linguísticas em amostras escritas no Brasil

Dias (2017) realizou um novo estudo com base nos estudos sociofuncionalistas sobre a indeterminação do sujeito, no entanto, os *corpora* utilizados foram da modalidade escrita do Projeto Para a História do Português Brasileiro (PHPB), bem como peças teatrais publicadas na Bahia no mesmo período. Em sua tese, ele investiga as principais estratégias de indeterminação do sujeito em textos escritos dos séculos XIX e XX na Bahia: Cartas de Leitores, Cartas de Redatores e Peças Teatrais. O quadro teórico-metodológico foi da Sociolinguística Variacionista. Nessa pesquisa, Dias (2017) buscou encontrar não só as formas canônicas de indeterminação do sujeito, sem sujeito lexicalmente preenchido ( $\emptyset+V3PP$ ,  $\emptyset+V+SE$ ,  $\emptyset+VINF$ ) mas também qualquer outro recurso utilizado pelos informantes para tal função, como aquelas frequentemente presentes em outros estudos: *você, a gente, nós, eles, voz passiva sem agente/VPSA*,  $\emptyset+V3PS$  e *sintagmas nominais* a exemplo de *o sujeito, o indivíduo e um homem*.

As variáveis foram divididas por Dias em três categorias: (i) contextos extralinguísticos (período de publicação das cartas/peças teatrais e o gênero textual); (ii) contextos linguísticos (flexão do verbo, tipo de oração, transitividade verbal, preenchimento do sujeito, estrutura do núcleo do predicado, concordância e posição do argumento interno do verbo, e a ausência x presença de preposição antes do verbo no infinitivo mais o *se*; e (iii) os contextos funcionais: função da indeterminação e o grau de indeterminação. As frequências dos recursos de indeterminação levantados no *corpus* de Dias foram as seguintes:  $\emptyset+V+SE$  (49,1%), *Nós* (20,1%),  $\emptyset+V3PP$  (12,9%). As demais variantes *SN*,  $\emptyset+VIN$   $\emptyset+V3PS$ , *a gente, eles, VPSA* e *você* não se mostraram significativas na amostra.

Dias (2017) encontrou estratégias parecidas daquelas verificadas em seu estudo de 2010, (Carvalho, 2010) no entanto, não foi registrada as estratégias *VPSA* e a pronominal *eu*. Semelhantemente, as formas pronominais *eles* e *você*, *VPSA* quase não apareceram, mas que também foi identificado por Carvalho (2010) de modo que o *você* foi uma das mais empregadas na fala da cidade de Salvador. Os resultados mostraram que a forma mais canônica ( $\emptyset+V+SE$ ) foi a que prevaleceu nos textos publicados entre nos séculos XIX e XX na Bahia, como também mostraram a inovação ao considerar a estratégia “nós”, para indeterminar o sujeito, sendo assim, a segunda mais utilizada, principalmente, quando se pretendia mostrar o *descomprometimento do agente* ou a *desfocalização do sujeito*.

Os gêneros epistolares foram os que mais apresentaram estratégias de indeterminação quando exerciam a função de *desfocalização do sujeito*, ou de uma *economia linguística* como também estavam mais presentes no XIX. O autor afirma que, como as peças teatrais se aproximam mais da oralidade, justifica-se o uso de utilizar mais estratégias e com mais frequências das formas mais inovadoras, no século XX.

As variáveis sociolinguísticas foram as mais favorecedoras para o emprego das estratégias de indeterminação, na seguinte ordem: o gênero textual com as estratégias ( $\emptyset+V+SE$ ,  $\emptyset+V3PP$ ,  $\emptyset+VINF$ ,  $\emptyset+V3PS$ , *VPSA*, nós e a gente). Em seguida, o período de publicação das cartas e peças teatrais ( $\emptyset+V+SE$ ,  $\emptyset+V3PS$ , nós, a gente e os sintagmas nominais). Sobre as variáveis funcionais, em lugar, o grau de indeterminação com as estratégias ( $\emptyset+V+SE$ ,  $\emptyset+V3PP$ ,  $\emptyset+V3PS$ , nós e a gente) e, em segundo lugar, *a função da indeterminação* com ( $\emptyset+V3PP$ , nós e sintagmas nominais). Sobre as variáveis linguísticas, *a flexão do verbo* apareceu com o ( $\emptyset+V+SE$ ,  $\emptyset+V3PP$ , *VPSA* e nós; o tipo de oração ( $\emptyset+V+SE$ ,  $\emptyset+VINF$  e os sintagmas nominais)

Dias (2017) reforça a importância de analisar as variáveis funcionais e seu trabalho mostrou bem isso. Segundo o autor, elas foram as mais selecionadas em todas as análises realizadas. Desse modo, discutir a função e o grau de indeterminação é imprescindível para compreender melhor as estratégias de indeterminação do sujeito.

#### 4.2.3.3 Estudos Linguísticas em amostras de fala em Angola

Teixeira e Almeida (2011) investigaram as estratégias ligadas ao uso do sujeito indeterminado comparando com o português de Angola e verificaram que, tanto no Brasil quanto em Angola, são utilizadas uma variedade de formas para indeterminar, e que, na norma vernacular, há o uso do *você* como uma forma preferida. Teixeira (2008) já tinha constatado

(em estudo anterior) que o *você* já era usado em situações mais íntimas ou universais pelos angolanos nas distintas faixas etárias e níveis de escolaridade, sobretudo no uso da norma popular pelos angolanos falantes das línguas nacionais. Segundo Teixeira e Almeida (2011, p. 112-113) com apoio em Teixeira (2008), o *você* já é bastante utilizado e esse fato pode demonstrar um encaixamento na estrutura social da mudança. Além da forma *você* já ser utilizado em situações reais como marca de segunda pessoa, passou também a ter um papel mais genérico, de indeterminado.

As autoras compararam com dados da segunda metade do século XIX por brasileiros com o objetivo de apresentar evidências de que o PA possa estar realmente passando por mudanças semelhantes ao PB e se afastando do PE, já atestados anteriormente por Galves (2008) e Teixeira (2008), pois o aumento do sujeito pleno e a entrada de *você(s)* e do *a gente* no paradigma pronominal do PB fez também refletir no uso de formas para indeterminação do sujeito. Processo parecido vem ocorrendo com o português falado em Angola.

As pesquisadoras utilizaram uma amostra de 20 informantes, 11 analfabetos ou informantes com pouco tempo de escolarização e 09 com ensino superior, concluído ou em curso e analisaram três fatores sociais: o sexo/gênero, a faixa etária e a língua nativa (português e línguas nacionais). Na pesquisa, os fatores linguísticos verificados foram a *presença ou ausência do pronome*, o *discurso diretivo X não-diretivo* e a *estrutura do predicado* (simples/complexa). As formas (nulas e plenas) consideradas foram *nós*, *você*, *tu*, *eles*, *a gente* e o clítico *se*.

Os resultados mostram que forma mais utilizada pelos falantes luandenses para indeterminar o sujeito foi a variante *nós* (38%), seguido do clítico *se* (17%) e por *eles* (16%) e, por último, a variante *você* (14%). Como os dados de P6, conforme as autoras não têm a característica de inclusão do falante, elas acreditam que a forma *você* e o clítico *se* concorrem entre si.

Com o cruzamento dos fatores escolaridade e língua nativa, Teixeira e Almeida (2011) encontraram os seguintes resultados para indeterminação do sujeito para a variante *você* pelos falantes das línguas nacionais: 30% para analfabetos e 12% para o nível superior. Já em relação aos falantes nativos, a frequência absoluta também foi maior entre os analfabetos (19%) e do nível superior (3%). A variante *nós* foi usada em maior quantidade pelos universitários (33%) do que aqueles com menos escolaridade (19%) falantes da segunda língua. As autoras explicam esses dados apontando a evidência a favor da hipótese da TLI e o papel da escola no processo de aquisição da norma culta. Outro ponto que reforça a hipótese da TLI apontada pelas autoras é, com base em (Lucchesi, 2003), o fato de *você* ter sido, primeiramente, utilizado pelos falantes

do Português como segunda língua, grande parte eram analfabetos e excluídos da sociedade. O *você*, ao ser usado pelo grupo dominador para indicar distância do grupo dominado, acabou sendo um modelo de uso linguístico acessível pelos angolanos.

Em relação ao *se*, todos os dois grupos apagaram, no entanto, houve uma maior frequência pelos falantes das línguas nacionais, 56%, e entre os falantes de português, 29%. Os resultados encontrados diferem daqueles de Duarte (2007), quando ela compara o uso do sujeito indeterminado no PE e no PB falado, pois o clítico *se* é a estratégia mais utilizada (38%) pelos portugueses europeus. Duarte (2007) reforça o fato de nunca existir a ausência do clítico *se* no PE em relação à indeterminação se comparada a variedade brasileira: “a ausência da estratégia com zero no PE e a baixíssima ocorrência do *se* no PB” (Duarte, 2007, p. 17). Teixeira e Almeida (2011) afirmam que no PA, o *se* muitas vezes é omitido, principalmente, quando há discurso diretivo, sem falar que foi apagado mais pelo grupo social composto por falantes das línguas nacionais.

Conforme Teixeira e Almeida (2011), a entrada do *você* e a do *a gente* têm começado a aparecer na fala de jovens angolanos pouco escolarizados, o que pode ser um indício de transição de uma mudança de “baixo para cima”. Esse fenômeno no PA “está levando a um preenchimento da posição de sujeito e ao uso dos mesmos como estratégia de indeterminação do sujeito”. Os resultados do PA têm se aproximado do PB e se afastado do PE. Diante disso, e do que já foi discutido anteriormente aqui nesse trabalho, é imprescindível a realização de novos estudos sobre a situação sociolinguística na comunidade angolana.

No quadro 8 a seguir, são apresentados os principais resultados dos estudos desenvolvidos no Brasil e em Angola, tanto em amostras de fala, em sua maioria, com o objetivo de visualizar melhor os fatores que mais favoreceram em relação às estratégias de indeterminação do sujeito em cada um deles.

**Quadro 8: Resumo do resultado de pesquisas desenvolvidas no Brasil e em Angola sobre o fenômeno da indeterminação do sujeito.**

	<b>Estratégias de indeterminação do sujeito mais favorecedoras</b>
<p><b>MILANEZ (1982)</b></p> <p>Recursos de indeterminação do sujeito”, utilizou o corpus de fala de 20 informantes de São Paulo, do Projeto NURC/SP e, ao utilizar a perspectiva gerativista transformacional.</p>	<p>- Sujeito lexical preenchido (62%), com o recurso <i>a gente</i> (Oc. 361);</p> <p>- Formas não pronominalizadas. <i>A gente</i>, foi seguida pela partícula <i>se</i> (Oc. 205) e <i>você</i> (Oc. 168)</p>

<p style="text-align: center;"><b>ROLLEMBERG (1991)</b></p> <p>Os pronomes pessoais e a indeterminação do sujeito na norma culta de Salvador. Utilizou os dados de fala de pessoas com ensino superior completo, coletados, ainda na década de 1970, pelo Projeto NURC/SSA</p>	<ul style="list-style-type: none"> <li>- Não houve registro das formas verbais despronominalizadas (<i>terceira pessoa do singular acompanhada do se, terceira pessoa do plural e verbo infinito impessoal</i>);</li> <li>- Sobre as formas pronominais, as mais presentes foram <i>você</i> com 35,54%, <i>nós</i> com 32,23%, <i>a gente</i>, 25,12% e <i>eles</i>, em menor quantidade, com 7,11% dos dados.</li> <li>- A forma <i>você</i> foi a mais utilizada pelos falantes mais jovens (Faixa 1) e cultos de Salvador, com prevalência na categoria de texto <i>DID</i>, e a forma <i>eles</i> foi a mais utilizada pelos falantes mais velhos (Faixa 3).</li> </ul>
<p style="text-align: center;"><b>MENON (1994)</b></p> <p>Pesquisa sociovariacionista em sua tese intitulada “Analyse sociolinguistique de l’indétermination du sujet dans le portugais parlé au Brésil, a partir des données du NURC/SP” utilizando um corpus de fala urbana de São Paulo com 68 entrevistas.</p>	<p>Em relação às variantes linguísticas:</p> <ul style="list-style-type: none"> <li>- O <i>a gente</i>, devido a um processo de gramaticalização, tornou-se um pronome genérico, depois indeterminado e, ainda utilizados como valor de 1ª pessoa, principalmente plural;</li> <li>- <i>A pessoa</i> passou a ser um substitutivo do <i>a gente</i> pelos falantes como se estivesse ocupando um espaço vazio;</li> <li>- O pronome <i>se</i> se caracteriza como um sujeito de pleno direito;</li> <li>- <i>Você</i> em concorrência com a forma <i>tu</i> tornou-se um pronome pessoal de segunda pessoa, além de portar a função de indeterminação do sujeito.</li> </ul> <p>Em relação às variantes extralinguísticas:</p> <ul style="list-style-type: none"> <li>- As mulheres usaram as formas inovadoras, enquanto os homens utilizaram as estratégias mais formais de indeterminação.</li> <li>- Os mais jovens usaram as formas <i>eu, eles e você</i>, enquanto os falantes mais velhos utilizavam mais o verbo na terceira pessoa do plural e a voz passiva sintética, enquanto o verbo na terceira pessoa do singular foi utilizado pelas três faixas.</li> </ul>
<p style="text-align: center;"><b>SETTI (1997)</b></p> <p>A amostra de Setti (1997) foi constituída de um total de 72 entrevistas, 24 entrevistas para cada uma das capitais analisadas, pertencentes o Projeto VARSUL.</p>	<ul style="list-style-type: none"> <li>- Os tempos e modos verbais foram favorecedores para os recursos de Indeterminação, com enfoque ao tempo presente do indicativo.</li> <li>- Sobre a variável escolaridade, os comportamentos em relação as homens e mulheres foram parecidos e mostraram um acerta neutralidade.</li> <li>- Os falantes menos escolarizados preenchiam mais o sujeito do que os mais escolarizados (tendência a utilizar o <math>\emptyset</math>+V3PP).</li> </ul>

	<ul style="list-style-type: none"> <li>- Em relação à variável sexo/gênero, as mulheres se mostraram mais conservadoras e os homens utilizaram as formas mais inovadoras.</li> <li>- Sobre às localidades, as capitais utilizaram as formas <math>\emptyset + V3PS</math>, <i>a gente</i> e <i>você</i> com maiores frequências relativas e, em Porto Alegre e Florianópolis houve a predominância da forma <i>tu</i>, exclusivamente. As formas menos frequentes foram <i>vocês</i> e <i>VPASSINT</i>.</li> </ul>
<p style="text-align: center;"><b>SANTANA (2006)</b></p> <p>Estratégias de indeterminação do sujeito no discurso oral de duas comunidades rurais afrodescendentes do semiárido baiano: Piabas e Bananal/Barra dos Negros, na região de Rio de Contas. A amostra foi constituída de 24 entrevistas.</p>	<ul style="list-style-type: none"> <li>- <i>A gente</i> (45%), <math>\emptyset + V3PS</math> (20%), os mais produtivos. <i>Eles</i> (10,5%), <i>nós</i> (9%), <i>você</i> (5%), <math>\emptyset + V3PP</math> (6%), e as demais, para valores inferiores;</li> <li>- 30% dos recursos de indeterminação foram de formas sem sujeito lexical e 70% dos pronomes pessoais, em especial, da forma <i>a gente</i>.</li> <li>- Tempo e modo verbal, com todas as variantes, o mais utilizado foi o presente do indicativo</li> <li>- Sobre os fatores extralinguísticos: as mulheres usaram mais <i>a gente</i>, enquanto os homens variavam mais os recursos utilizando <i>você</i>, <math>\emptyset + V3PS</math>, <math>\emptyset + VINFINF</math>; faixa etária, as faixas 2 e 3 favoreceram mais o uso das formas indeterminadas, com destaque para <i>a gente</i> e <i>você</i> (faixa 2) e <math>\emptyset + V3PS</math> (faixa 3); Localidade: Piabas- comunidade mais isolada- favoreceu as formas mais conservadoras (<i>nós</i> e <math>\emptyset + V3PP</math>), e Bananal/Barra dos Negros utilizou a forma mais inovadora <i>a gente</i>.</li> </ul>
<p style="text-align: center;"><b>PONTE (2008)</b></p> <p>Estratégias de indeterminação do sujeito do português popular o interior do estado da Bahia com amostras de pesquisas realizadas nas comunidades de Barra e Bananal (Rio de Contas), Sapé (Valença), Helvécia (Nova Viçosa) e Cinzento (Planalto) e nas zonas rural e urbana de Santo Antônio de Jesus. Foi utilizado o <i>Corpus</i> Base do Português Afro-Brasileiro do Estado da Bahia além de 24 entrevistas coletadas em Santo Antônio de Jesus do Acervo de Fala Vernácula do Português Afro-Brasileiro do Estado da Bahia.</p>	<ul style="list-style-type: none"> <li>- <i>A gente</i> obteve 47% e <math>\emptyset + V3PS</math>, com 25%, mais produtivos. <i>Você</i> resultou em 9% e as demais variantes, com valores abaixo de 6%;</li> <li>- No nível genérico, prevaleceram as formas <i>a gente</i>, <i>você</i> e <math>\emptyset + V3PS</math>;</li> <li>- Nos níveis parcialmente definido e eu ampliado, destacaram-se os pronomes P4 (<i>a gente</i> e <i>nós</i>) e a forma <math>\emptyset + V3PS</math>.</li> <li>- No nível do agente específico indefinido, predominou a P6 (<i>eles</i> e <math>\emptyset + V3PP</math>);</li> <li>- A estratégia <math>\emptyset + V3PS</math>, teve maior relevância nos 4 níveis de referencialidade;</li> <li>- Em relação aos contextos linguísticos, a forma <i>antecedente</i> foi a única selecionadas;</li> <li>- Sobre os contextos sociais, a localidade do informante mostrou um <i>continuum</i> linguístico da forma <math>\emptyset + V3PS</math>;</li> <li>- <i>A gente</i>, <math>\emptyset + V3PP</math> e <math>\emptyset + V3PS</math> foram as escolhas dos falantes fora da comunidade;</li> <li>- Os falantes mais escolarizados, bem como as mulheres favoreceram a estratégia do pronome <i>nós</i>;</li> </ul>

	<p>- Ø + V3PS + SE, quase não foi usada pelos informantes.</p>
<p><b>CARVALHO (2010)</b></p> <p>Estratégias de indeterminação do sujeito utilizando 44 inquiridos, 32 pertencentes ao Programa de Estudos sobre o Português Popular Falado de Salvador (PEPP) e 12 do Projeto NURC/Salvador.</p>	<p>- As formas <i>você, a gente</i> e as <i>formas nominais</i> foram as mais presentes nos corpora;</p> <p>- Sobre as variáveis linguísticas, as <i>formas pronominais</i>, a exemplo de <i>você, a gente, nós, eles</i> e <i>eu</i> prevaleceram em relação as <i>não pronominais</i>, motivada, principalmente, pela variável <i>forma antecedente</i>, sobretudo quando houve a manutenção do referente;</p> <p>O Ensino Superior contribuiu: quanto maior escolaridade do falante, maior foi o leque de opções linguísticas.</p>
<p><b>TEIXEIRA E ALMEIDA (2011)</b></p> <p>Estratégias ligadas ao uso do sujeito indeterminado comparando com o português de Angola utilizando uma amostra de 20 entrevistas.</p>	<p>- Uso do <i>você</i> como uma forma preferida (utilizado em situações reais como marca de segunda pessoa e de papel mais genérico, indeterminador.</p> <p>- <i>nós</i> (38%), seguido do clítico <i>se</i> (17%) e por <i>eles</i> (16%) e, por último, a variante <i>você</i> (14%).</p> <p>- Em relação à escolaridade X língua, a forma <i>você</i> foi utilizada por falantes das línguas nacionais: analfabetos (30%), nível superior (12%).</p> <p>- Em relação aos falantes nativos, analfabetos (19%) e nível superior (3%).</p> <p>A variante <i>nós</i> foi usada em maior quantidade pelos universitários (33%) e aqueles com menos escolaridade (19%) com falantes da segunda língua.</p> <p>- Em relação ao <i>se</i>, todos apagaram, havendo maior frequência no uso dos falantes das línguas nacionais, 56%, e entre os falantes de português, 29%.</p>
<p><b>ASSUNÇÃO (2012)</b></p> <p>Entrevistas realizadas do tipo DID entre (2009 e 2012) do projeto “A língua portuguesa falada no semiárido baiano” da UEFS ao analisar a fala de 24 feirenses</p>	<p>- A variante <i>tu</i> não foi encontrada no <i>corpus</i> feirense com referência indeterminadora;</p> <p>- As formas consideradas padrão Ø + VERBO + SE e ØV3PP não foram significativas, sendo que, os poucos dados de ØV3PP encontrados foram realizados pelos falantes menos escolarizados;</p> <p>- Sobre a variante <i>você</i>, as variáveis linguísticas mais significativas foram o tipo de oração, tipo de verbo, preenchimento do sujeito e o tempo e o modo verbais (futuro do subjuntivo com P.R. 0.82, infinitivo, P.R. 0.76 e presente do indicativo, 0.56);</p> <p>- Em relação aos fatores extralinguísticos, todos os grupos foram selecionados na seguinte ordem – faixa etária (os mais jovens, P.R. 0.66), sexo/gênero (homens, P.R. 0.70) e escolaridade (ensino superior, P.R. 0.62).</p>

	<p>- <i>Você</i> foi a variante mais inovadora utilizada pelos jovens e mostraram uma mudança em curso.</p>
<b>SOUZA (2014)</b>	
<p>Formas de indeterminação do sujeito na fala de moradores da zona rural da Matinha (Feira de Santana) e de Casinhas (Jeremoabo). utilizou 24 inquéritos do tipo DID, 12 para cada localidade, gravados e transcritos pelo projeto “A língua portuguesa no semiárido baiano”, com sede no NELP/UEFS</p>	<p>- O pronome <i>a gente</i> (41%) e o <math>\emptyset+V3PS</math> (28%) foram os mais utilizados pelos falantes nas duas comunidades.</p> <p>- As variantes mais tradicionais, o <math>\emptyset+V+SE</math> e <math>\emptyset+V3PP</math> as ocorrências foram ínfimas.</p> <p>- A variável <i>formas nominais</i> (17%) foi a terceira estratégia mais utilizada.</p> <p>- Variáveis linguísticas: o tempo/modo verbal e o paralelismo foram mais significativos; a forma <i>a gente</i> (favorecida com o presente do indicativo, o futuro do presente do indicativo), e após forma idêntica. Houve favorecimento diante de verbos intransitivos e, em orações subordinadas.</p> <p>- Variáveis sociais: grau de escolaridade, o sexo e a localidade dos informantes se sobressaíram, sendo que o <i>a gente</i> foi o mais utilizado por mulheres, da faixa 3 e que possuíam 2 a 3 anos de escolarização.</p>
<b>MENDONÇA E NASCIMENTO (2015)</b>	
<p>Analisa o fenômeno da indeterminação do sujeito ao analisar a amostra Rede Social de Informantes Universitários de Itabaiana/SE (Araújo; Santos; Freitag, 2014)</p>	<p>- As estratégias nominais foram as maiores favorecedoras, com 71% dos dados, seguidos pelas sintáticas e pelas formas nominais com 25% e 4%, respectivamente.</p> <p>Sobre a variável sexo/gênero foi característico de preferência/escolha por algumas das formas elencadas, a exemplo do sexo feminino com a variante <i>a gente</i>.</p>
<b>DIAS (2017)</b>	
<p>Estudo sociofuncionalista sobre a indeterminação do sujeito, utilizando corpora da modalidade escrita do Projeto Para a História do Português Brasileiro (PHPB), e peças teatrais publicadas na Bahia no mesmo período. Investiga as principais estratégias de indeterminação do sujeito em textos escritos dos séculos XIX e XX na Bahia: Cartas de Leitores, Cartas de Redatores e Peças Teatrais</p>	<p>- As frequências dos recursos de indeterminação foram: <math>\emptyset+V+SE</math> (49,1%), <i>Nós</i> (20,1%), <math>\emptyset+V3PP</math> (12,9%). As demais variantes <i>SN</i>, <math>\emptyset+VIN</math> <math>\emptyset+V3PS</math>, <i>a gente</i>, <i>eles</i>, <i>VPSA</i> e <i>você</i> não se mostraram significativas na amostra.</p> <p>- A forma mais canônica (<math>\emptyset+V+SE</math>) prevaleceu nos textos publicados como também mostraram a inovação do uso do “nós”, para indeterminar o sujeito, sendo assim, a segunda mais utilizada, principalmente, quando se pretendia mostrar o <i>descomprometimento do agente</i> ou a <i>desfocalização do sujeito</i>.</p> <p>- Os gêneros epistolares foram os que mais apresentaram diferentes estratégias quando exerciam a função de <i>desfocalização do sujeito</i>, ou de uma <i>economia linguística</i>.</p> <p>- As variáveis sociolinguísticas foram as mais favorecedoras: o gênero textual (<math>\emptyset+V+SE</math>, <math>\emptyset+V3PP</math>, <math>\emptyset+VIN</math>, <math>\emptyset+V3PS</math>, <i>VPSA</i>, <i>nós</i> e <i>a gente</i>). Em seguida, o período de publicação das cartas e</p>

	<p>peças teatrais (<i>Ø+V+SE, Ø+V3PS, nós, a gente e os sintagmas nominais</i>).</p> <ul style="list-style-type: none"> <li>- Sobre as variáveis funcionais: o grau de indeterminação com as estratégias (<i>Ø+V+SE, Ø+V3PP, Ø+V3PS, nós e a gente</i>) e a função da indeterminação com (<i>Ø+V3PP, nós e sintagmas nominais</i>).</li> <li>- Sobre as variáveis linguísticas, a flexão do verbo apareceu com (<i>Ø+V+SE, Ø+V3PP, VPSA e nós</i>; o tipo de oração (<i>Ø+V+SE, Ø+VINFINF e os sintagmas nominais</i>))</li> </ul>
<p style="text-align: center;"><b>GUIMARÃES (2018)</b></p> <p>Pesquisa geosociolinguística para identificar as formas de indeterminação mais recorrentes em 7 cidades baianas das sete mesorregiões do estado. O <i>corpus</i> utilizado foi composto de 28 entrevistas do projeto Atlas Linguístico do Brasil (ALiB), ao utilizar formato perguntas pré-elaboradas baseadas no questionário fonético-fonológico e semântico-lexical.</p>	<ul style="list-style-type: none"> <li>- O Grupo 1 (<i>gramaticais e suas variações</i>) apresentou maior percentual (49,7%), seguido das <i>nominais e pronominais</i> com 32,7% e 17,6%, respectivamente. Nesse grupo, a faixa etária foi selecionada (maior favorecimento dos falantes da faixa etária 2 (verbo na 3ª pessoa do singular sem 'se') e desfavorecimento entre falantes da faixa etária 1 (preferência por verbos na 3ª pessoa do plural).</li> <li>- A forma <i>3ª pessoa do singular com 'se'</i> apresentou valores acima de 65% em todas as cidades analisadas.</li> <li>- Sobre as variantes do Grupo 2, o subgrupo mais favorecedor foi das variantes cuja estrutura morfossintática era de um artigo acompanhado de substantivo ou pronome indefinido (55,8%)</li> <li>- As cidades/mesorregiões Vitória da Conquista, Barreiras e Barra – parte oeste do estado - favoreceu as formas nominais do grupo A e as demais desfavorecendo.</li> <li>- A forma <i>a gente</i> alcançou 63% da amostra analisada, ultrapassando a forma <i>nós</i>, com exceção de Barreiras (36,4% dos dados, apenas)</li> </ul>

Na Seção 5, a seguir, são discutidos os pressupostos teóricos metodológicos que são tomados para o estudo da indeterminação do sujeito na comunidade de Angola.

## 5 PRESSUPOSTOS TEÓRICO-METODOLÓGICOS

[...] Estudos mais detalhados do contexto social em que a língua é usada mostram que muitos elementos da estrutura linguística estão envolvidos na variação sistemática que reflete tanto a mudança temporal quanto os processos sociais extralinguísticos. (Labov, 2008 [1972], p. 140)

Nesta subseção, são descritos os pressupostos teórico-metodológicos da Teoria da Variação e Mudança bem como o *corpus* que foi utilizado do Projeto “Em busca das raízes do Português Brasileiro”, sediado no NELP. Apresentam-se informações sobre a comunidade linguística investigada (Luanda), seus informantes, a maneira como as entrevistas foram realizadas e transcritas, o método quantitativo de análise e os instrumentos utilizados a fim de possibilitar uma análise, que será apresentada na Seção 6. Também são apresentadas as variáveis controladas nessa pesquisa, as dependentes e independentes (linguísticas e sócio-históricas) com base em estudos prévios sobre o fenômeno da indeterminação do sujeito bem como de peculiaridades do próprio *corpus* analisado.

### 5.1 SOCIOLINGUÍSTICA LABOVIANA: TEORIA DA VARIAÇÃO E MUDANÇA LINGUÍSTICA

Alguns dos estudos que antecederam a Sociolinguística Variacionista pertenciam a correntes de conhecimento que viam a língua como um sistema homogêneo e estático, separados de uma comunidade de fala, como é o caso das correntes estruturalistas e gerativistas. A primeira, com base nos estudos de Saussure, tem a língua como um objeto sincrônico, ocupando-se estritamente de um estado de língua imobilizada fora de seu tempo, sendo a homogeneidade uma característica intrínseca à língua. Saussure, por uma escolha metodológica, elege a sincronia, retirando a dimensão histórica da língua (processo de mudança) na caracterização da ciência. A segunda, a corrente gerativista, contou com o principal nome de Noam Chomsky, que buscava, naquele momento, colocou em evidência o falante ideal, a sua competência e não os locutores reais em seu uso concreto da linguagem. Assim, os estudos se pautaram mais no conhecimento interno das regras que comandavam a língua do que nas manifestações da língua em uso real nas diversas interações sociocomunicativas dos falantes.

Diante de um quadro como o delineado acima, a teoria sociolinguística surge na década de 1960, provocando uma renovação na base dos estudos linguísticos. O intuito, diferentemente

das correntes estruturalistas e gerativistas, era estudar as relações que estão imbricadas entre o homem e a língua e o comportamento linguístico resultante dessa inter-relação nas diferentes comunidades. A língua passa a ser vista como um sistema heterogêneo utilizada pelos falantes no processo de comunicação e passível de descrições e de análises científicas.

Weinreich, Labov e Herzog (2006 [1968]) comprovam que as línguas variam motivadas por fatores internos e externos ao sistema. O condicionamento motiva a variação e esse mesmo condicionamento garante o entendimento entre os falantes em uma situação de variação. Desta forma, os autores apresentam uma teoria de mudança pautada na explicação de como a estrutura linguística se transforma ao longo do tempo sem que gere caos enquanto mudanças estão acontecendo em uma dada comunidade. Fica evidente que a heterogeneidade linguística passa a ser o centro dos estudos da variação e/ou da mudança da língua em que a coexistência de diversos falares não acarreta prejuízo de entendimento pelos falantes, comprovando assim uma sistematização no aparente “caos linguístico” dos fatores condicionadores do processo variável oriundo das pressões sociais.

Considerando fatores internos e externos à língua, a Teoria da Variação e Mudança Linguística ou Sociolinguística Variacionista criou uma metodologia própria e segura, capaz de fornecer explicações e descrições de maneira sistemática dos fenômenos variáveis encontrados nas comunidades de fala sincrônica ou diacronicamente, e sempre relacionando aos contextos sociais e linguísticos. Sobre a mudança linguística apontada por Weinreich, Labov e Herzog (2006 [1968]), há cinco problemas que precisam ser levados em conta e que podem nortear o estudo: condicionamentos, encaixamento, transição, implementação e avaliação.

O problema dos *Condicionamentos* ou restrição busca compreender o conjunto de mudanças possíveis e as condições que podem levar a uma mudança ao longo do tempo. A Restrição está ligada ao conjunto de possíveis mudanças e condições que podem acontecer numa determinada estrutura. Desse modo, deve-se considerar as motivações para os usos de determinadas variantes. Nesta pesquisa, em andamento, buscou-se investigar os usos mais frequentes e menos frequentes para as variantes tanto inovadores quanto aquelas cristalizadas pelas gramáticas normativas no entendimento das estratégias de indeterminação do sujeito pela comunidade analisada. Vale ressaltar que, como a variação é uma propriedade regular do sistema, e não acontece aleatoriamente, os resultados são mais coerentes, fidedignos se ambos os fatores linguísticos e extralinguísticos forem analisados, conjuntamente.

O segundo problema, a *Transição*, diz respeito à compreensão de como acontece a variação ou mudança de um estágio para o outro. Para Weinreich, Labov e Herzog (2006 [1968], p.122), “[...] a mudança se dá (1) à medida que o falante aprende uma forma alternativa,

(2) durante o tempo em que as duas formas existem em contato dentro de sua competência, e (3) quando uma das formas se torna obsoleta”. Assim, fica claro que as formas coexistem e suas variantes estão em concorrência até o ponto que uma delas vai perdendo força em relação à outra e se torna aceita pela comunidade como um todo. A mudança linguística não acontece de forma abrupta, ela é lenta e as formas antigas não são substituídas de uma para outra pois existe uma fase intermediária (coexistência das variantes) o que não implica em desordem do sistema. A heterogeneidade da língua continua sistemática e ordenada por todo o período de transição, muitas vezes imperceptível até mesmo para os falantes.

O *Encaixamento* é um terceiro problema que a teoria laboviana busca entender. Nesse sentido, é muito importante fazer uma análise linguística complementando com a observação de sua interação com a estrutura social da comunidade de fala. Sobre o primeiro, o encaixamento na estrutura linguística, se entende como a capacidade e o controle acontecem pelos membros da comunidade de fala ao inserir a forma inovadora na estrutura da língua atribuindo-lhe um valor com a sua funcionalidade. Sobre o segundo, o encaixamento na estrutura social, Weinreich, Labov, Herzog (2006 [1968], p. 123) afirmam que “a estrutura linguística mutante está ela mesma encaixada no contexto mais amplo da comunidade de fala, de tal modo que variações sociais e geográficas são elementos intrínsecos da estrutura”. Para os teóricos supracitados, os fenômenos que se encontram em processo de mudança se encaixam no sistema abstrato de forma que ele continua estruturado e não acarreta prejuízo aos falantes. Desse modo, mesmo uma mudança se encaixando em outras matrizes de mudanças, tudo na língua ainda é bem ordenado e estruturado

O quarto problema, a *Avaliação* diz respeito à atitude subjetiva e consciente dos falantes sobre as formas da língua em variação ou mudança. Sobre esse problema, Labov (2008 [1972], p. 152) afirma que a adoção de novas formas linguísticas depende do nível de consciência do falante e podem ocorrer em duas direções: mudanças vindas de baixo (change from below), quando uma variante inovadora aparece inicialmente no vernáculo sem a consciência de seus falantes e se estende para outros estratos sociais mais altos, e as mudanças vindas de cima (change from above), que são introduzidas pela classe dominante e de maior prestígio, e refletem um dialeto adquirido depois da aquisição do vernáculo. Como os falantes apresentam um nível relativamente alto de consciência social, as mudanças aparecem primeiro em um estilo de fala mais monitorada, mais cuidada. Para Labov (2008 [1972], p. 152), as mudanças que são conduzidas por forças que vem de baixo estão “[...] abaixo do nível da percepção consciente [...]” e acrescenta que “[...] as pressões vindas de baixo operam sobre sistemas linguísticos

inteiros, em resposta a motivações sociais que são relativamente obscuras e mesmo assim têm a maior importância para a evolução geral da língua”.

A *Implementação*, o quinto problema, se propõe a entender as diferenças na aplicação das regras variáveis e a propagação da mudança, pois, em algum momento, um fenômeno deixa de ser variável e passa a ser considerado parte de uma estrutura linguística de uma determinada comunidade. Nesse problema, são discutidas as razões pelas quais uma mudança acontece em determinada comunidade e não em outra, bem como ela acontece em um período e não em outro. Para que as perguntas que o linguista faz sejam respondidas, suas respostas só são possíveis se todos os outros 4 problemas forem também considerados. É a combinação de elementos investigados, de diferentes questões que possibilitam as respostas sobre o fenômeno e sobre a comunidade analisada.

Nesta tese, o problema empírico da restrição/dos fatores condicionantes tem o seu destaque, visto que há uma busca pela compreensão das motivações internas e externas através do controle dos fatores sócio-históricos quanto às estratégias de indeterminação do sujeito na fala de luandenses. Deste modo, a variação só pode ser mais bem compreendida se ambos os fatores, linguísticos e extralinguísticos, forem partes de estudo.

Para uma melhor descrição do fenômeno linguístico, muitas vezes faz-se necessária uma análise que combine estudos sincrônicos e diacrônicos, ou seja, amostras de tempo aparente e de tempo real. O estudo da sincronia, estudo em tempo aparente, compreende uma análise das amostras de informantes de diferentes faixas etárias em um mesmo momento histórico, assim, são feitas análises com o objetivo de perceber indícios de mudanças em progresso ou até de mudanças já implementadas. Assim, esta pesquisa caracteriza-se como um estudo de tempo aparente, pois buscou analisar, além do comportamento de grupos etários falantes do português de Luanda, outras variáveis sócio-históricas, como o sexo, a escolaridade, a língua materna e o local de nascimento dos falantes, em conjunto com os fatores intralinguísticos como o tempo/modo verbal, o tipo de verbo, o paralelismo, o grau de Indeterminação, o tipo de oração, entre outros.

Labov (2008 [1972]) afirma que o estudo da língua deve ser feito observando-se o seu contexto linguístico e extralinguístico e, segundo o autor,

[...] os procedimentos da linguística descritiva se baseiam no entendimento de que a língua é um conjunto estruturado de normas sociais. [...] Estudos mais detalhados do contexto social em que a língua é usada mostram que muitos elementos da estrutura linguística estão envolvidos na variação sistemática que reflete tanto a mudança temporal quanto os processos sociais extralinguísticos. (Labov, 2008 [1972], p. 140)

Diante disso, é papel da sociolinguística descrever a relação existente entre os fatos sociais, aqueles que são exteriores à língua, como idade, sexo/gênero, escolarização, classe social, profissão, registro/estilo, entre outros, e os fatos internos à língua. Sobre as variáveis sociais utilizadas, é importante analisar, além da faixa etária, o sexo, pois homens e mulheres podem apresentar diferentes papéis sociais e, conseqüentemente, comportamentos distintos em relação ao processo de mudança. As mulheres podem fazer o uso das variantes mais conservadoras e de prestígio, pois elas costumam empregar as normas gramaticais prescritas como da língua-padrão. O inverso também pode ser considerado, visto a existência de estudos que mostram que as mulheres também podem encabeçar a mudança. Não somente, a escolaridade é uma variável importante, pois a escola atua como agente de manutenção das formas de prestígio e estigmatiza os diferentes falares inovadores como erros. Assim, o nível de escolaridade do falante pode corroborar ou não a presença das formas daquelas aceitas pelas gramáticas tradicionais.

Conforme Teixeira e Almeida (2011), o contato é um fator forte a ser considerado no processo de aquisição do português como segunda língua, principalmente por africanos, o que tem levado pesquisadores a estudarem o a variedade do Português de Angola sem deixar de lado o processo de transmissão linguística irregular no território africano, algo que também permeou a formação sócio-histórica na variedade brasileira. Assim, encontrar respostas para as perguntas já apontadas podem auxiliar na compreensão se há evidências de que as mudanças ocorridas com o PB no século XIX (Tarallo, 2018 [1993], p. 79) são semelhantes aos processos atuais do PA, além de permitir um estreito diálogo sobre o papel do contato linguístico na formação das variedades nacionais.

Desse modo, neste trabalho buscou-se uma melhor compreensão da situação sociolinguística de Luanda através da análise do fenômeno morfossintático da indeterminação do sujeito sem perder de vista o cenário marcado por intensas relações sociais dos africanos dessa comunidade de fala com o povo português ao longo de 400 anos.

Na subseção seguinte, descreve-se o método da Sociolinguística Laboviana, bem como a apresentação da ferramenta computacional que possibilita um estudo quantitativo dos dados. Esses resultados estatísticos permitem que o pesquisador faça uma análise mais acertada do fenômeno variável sob o enfoque teórico-metodológico adotado.

## 5.2 O MÉTODO DA SOCIOLINGUÍSTICA QUANTITATIVA LABOVIANA

Na realização desta pesquisa, utilizou-se o modelo teórico-metodológico da Teoria da Variação e Mudança (LABOV, 2008[1972]), corrente teórica que considera a heterogeneidade sistemática inerente à língua e que possibilita um tratamento eficaz e estatístico dos dados por meio de ferramentas computacionais.

Para Naro (2010, p. 25), a ferramenta que a metodologia da Teoria da Variação e Mudança utiliza é muito poderosa e segura. Além disso, ela pode ser usada para estudos de vários fenômenos linguísticos, em seus diversos níveis, embora, ele alerta para as limitações afirmando que:

São as do próprio linguista, a quem cabe a responsabilidade de descobrir quais são os fatores relevantes, de levantar e codificar os dados empíricos corretamente e, sobretudo, de interpretar os resultados numéricos dentro de uma visão teórica da língua. O progresso da ciência linguística não está nos números em si, mas no que a análise dos números pode trazer para nosso entendimento das línguas humanas. (Naro, 2010, p. 25)

Considerando o caráter heterogêneo, variável e, ao mesmo tempo, ordenado e sistemático da língua, na tentativa de identificar as condições que fazem com que os falantes usem certas estruturas em detrimento de outras, a sociolinguística quantitativa lança uso de uma pesquisa com bases empíricas ao utilizar um modelo de análises estatísticas para testagem de hipóteses.

William Labov, em 1969, foi quem criou o primeiro modelo estatístico, um modelo matemático consistente para o tratamento dos dados linguísticos, um modelo aditivo em que  $f$  considerava a soma dos fatores contextuais, no entanto, posteriormente, foi abandonado pois dava muitos problemas de ordem técnica.

Em 1974, outro modelo, com características multiplicativas para o tratamento dos dados foi criado por Henrietta Cedergren e David Sankoff, mas, nem o modelo anterior de Labov, nem esse último foram capazes de suprir satisfatoriamente a pesquisa na época. Conforme Naro (2010, p. 20),

Cedergren e David Sankoff propuseram uma interpretação probabilística que substituiu frequências por probabilidades, e apresentaram como função básica de atuação conjunta de fatores o modelo clássico de efeito simultâneo de fatores independentes.

Em 1978, surgiu outro modelo, um modelo de regressão logística de Pascale Rousseau e David Sankoff que foi possível atender as necessidades que os modelos anteriores não supriram. Desse modo, houve o aperfeiçoamento dos modelos estatísticos até chegar ao modelo logístico

aplicado atualmente nas pesquisas sociolinguísticas. Chegou-se à descrição de uma curva com um formato em S<sup>46</sup> que se aplica bem à descrição da variação sincrônica, visto que esse tipo de variação corresponde a um recorte de um processo de mudança. Com o auxílio do computador, tornou-se mais viável a utilização desse modelo logístico pois a quantidade de cálculos era muito grande. Nesse contexto, o pacote de Programas VARBRUL (*Variable Rules*) contribuiu com uma análise mais rápida e com confiabilidade nos resultados. Assim, constituiu-se em um poderoso instrumento de auxílio na análise de fenômenos linguísticos variáveis.

Assim, com base na Sociolinguística Laboviana, esta pesquisa objetiva identificar os contextos linguísticos e sociais que podem atuar nas estratégias de indeterminação do sujeito pelos residentes em Luanda. As variáveis, ou grupos de fatores, foram, previamente, elencados a partir de estudos já realizados e, que analisaram majoritariamente dados de língua falada sobre a indeterminação do sujeito, conforme apresentados na Seção 3.

### 5.2.1 A ferramenta computacional *Goldvarb X*

Como descrito na subseção anterior, o pacote VARBRUL, um pacote para análise dos dados linguísticos variáveis através do tratamento estatístico e probabilístico estava de acordo com a abordagem variacionista. O VARBRUL permite, assim, a análise de regras variáveis e fornece cálculos de frequência, percentuais e pesos relativos ao associar cada fator das variáveis independentes com a aplicação da regra.

Atualmente, esse programa possui uma versão mais atualizada, o *GOLDVARB X* e está disponível gratuitamente. Segundo Guy e Zilles (2007, p. 73), “[...] a realização de análises quantitativas possibilita o estudo da variação lingüística, permitindo ao pesquisador apreender sua sistematicidade, seu encaixamento lingüístico e social e sua eventual relação com a mudança lingüística”.

Com o *Goldvarb X*, podem ser feitas análises univariadas, também chamadas de unidimensionais, análises multivariadas ou multidimensionais e as tabulações cruzadas. As análises univariadas são aquelas que testam o efeito de uma variável independente sobre uma variável dependente e cujos resultados são dados em forma de frequências absolutas ou relativas. Já, as análises multivariadas permitem a investigação de situações em que a variável linguística influenciada por múltiplas variáveis independentes. Ainda sobre as análises

---

<sup>46</sup> Conforme Naro “quando temos duas alternativas em competição e uma possui alguma vantagem que torna seu uso futuro mais provável, a forma pela qual se dá a distribuição no tempo é uma curva com o formato em S deitado. (NARO, 2010, p. 22)

multivariadas, há a investigação dos efeitos, o valor da significância estatísticas das variáveis independentes com a dependente.

Em relação aos tratamentos dos dados nas pesquisas, o *Goldvarb X* faz a análise de pesos relativos de duas variantes (*binominal*), em um nível (*one level*) ou em múltiplos níveis (*up and down*). A partir do momento em que o programa começa a fazer a seleção, ele apresenta o *input* inicial, que corresponde ao percentual total da regra de aplicação, o nível geral de uso de um determinado valor da variável de aplicação da regra. O *input* final é apresentado, posteriormente, pelo programa para mostrar que ele corrigiu algum desequilíbrio na amostra.

O programa oferece um nível de significância estatística de 0.5, que é a probabilidade de se observar determinada distribuição em caso de a hipótese nula ser verdadeira, sendo que valores acima de 0.5 são favorecedores ao uso de determinada forma e abaixo desfavorecedoras. Já o *log likelihood* é um logaritmo de verossimilhança responsável por medir a aproximação entre o modelo e os dados observados.

Após os dados serem codificados pelo pesquisador, eles devem ser submetidos ao programa para fazer a seleção estatística dos grupos de fatores por ordem de relevância, emitir os resultados estatísticos e probabilísticos, e fornecer valores confiáveis com os pesos relativos (PR) que indicam o efeito de cada um dos fatores sobre a variável dependente, que é o fenômeno linguístico analisado.

É importante ressaltar que cabe ao pesquisador a verificação dos seus dados, a análise (quantitativa e qualitativa) com a ajuda do arcabouço da teoria da Sociolinguística laboviana. Dados estatísticos são apenas indicadores empíricos, estatísticas inferenciais adicionais que são utilizados na busca de respostas sobre os fenômenos variáveis.

### 5.2.2 Etapas da pesquisa

Muitas etapas foram seguidas até aqui. Os tópicos que seguem apresentam os passos seguidos nessa pesquisa.

O tópico (i) corresponde apenas às etapas pré pesquisa, enquanto as demais a saber:

(i) Preparação do Anteprojeto da pesquisa e arguição pela banca de seleção como etapa obrigatória e eliminatória (etapa de seleção para aprovação na seleção do doutorado em caráter de aluno regular); preparação do projeto de pesquisa nas aulas de Metodologia da Pesquisa, tendo como base o anteprojeto; defesa pública do projeto de pesquisa; ajustes do projeto pós-

defesa com vistas ao refinamento da proposta da pesquisa de doutoramento e à submissão do projeto no Comitê de Ética em Pesquisa em Seres Humanos - CEP/UEFS (foram apresentados ao CEP a autorização da instituição proponente (UEFS) através do orientador e da chefia do Departamento de Letras da UEFS, o termo de Anuência da coordenação do NELP, responsável pelo material a ser utilizado, o projeto “Em busca das raízes do Português Brasileiro” e o parecer de aprovação já emitido na época que os dados foram coletados, a dispensa dos Termos de Consentimento Livre Esclarecido para os participantes da pesquisa, entre outros documentos requisitados para a aprovação pelo Comitê de Ética da UEFS;

(ii) Leitura bibliográfica de estudos já realizados sobre o fenômeno da indeterminação do sujeito em território brasileiro e angolano, além de estudos dos arcabouços teóricos e metodológicos;

(iii) Acesso as entrevistas, leitura do material e triagem dos participantes;

(iv) Checagem das entrevistas (áudios e transcrições grafemáticas) anteriormente realizadas pelos colaboradores do Projeto “Em busca das raízes do Português Brasileiro”;

(v) Transcrição e leitura das entrevistas que ainda não foram realizadas com a ajuda dos orientandos de Iniciação Científica da graduação em Letras da profa. Dra. Silvana Silva de Farias Araújo (coordenadora do Projeto sediado no NELP);

(vi) Controle das variáveis analisadas com base em pesquisas prévias e da situação sociolinguística da comunidade investigada;

(vii) Exame de qualificação;

Após o exame de qualificação, outras ações foram realizadas para o refinamento da tese e consistiram na quantificação dos dados, nos ajustes do texto com base nas sugestões feitas pela banca e orientadora, na ampliação da leitura bibliográfica e na intensificação da escrita final para a defesa pública desta tese, mais bem discriminadas a seguir:

(i) Adequação das sugestões e correções sugeridas pela banca avaliadora da qualificação;

(ii) Releitura das entrevistas selecionadas com vistas à identificação do fenômeno da indeterminação do sujeito nas entrevistas;

(iii) Codificação das ocorrências de acordo com as variáveis controladas, definição das variáveis independentes, seus possíveis fatores condicionadores em relação a variável dependente a ser testada nesta pesquisa;

(iv) Submissão dos dados codificados ao *Goldvarb X*;

(v) Análise quantitativa e qualitativa dos dados seguindo os pressupostos da Teoria da Variação e Mudança Linguística;

- (vi) Ajustes da escrita, correções necessárias e aprofundamento do material;
- (vii) Defesa da tese e entrega final ao PPGEL/UEFS.

Na próxima subseção são apresentados como o *corpus* foi constituído, o Projeto “Em busca das raízes do Português Brasileiro”, as entrevistas e as transcrições realizadas. A comunidade de Luanda e os participantes da pesquisa são apresentados também. São expostos as amostras, o perfil dos participantes e as variáveis controladas - as dependentes e os condicionamentos de natureza externa (extralinguísticas) e os intralinguísticos que possibilitam uma melhor explicação do fenômeno aqui investigado.

### 5.3 O *CORPUS*

Os dados investigados nesta pesquisa foram extraídos de uma amostra constituída de 32 entrevistas sociolinguísticas do tipo Diálogo entre Informante e Documentador (DID), gravadas com falantes do português na cidade de Luanda-Angola, nos anos de 2008 e 2013. Essas entrevistas fazem parte do acervo do projeto “Em busca das raízes do Português Brasileiro”, pertencente ao Núcleo de Estudos em Língua Portuguesa do Departamento de Letras e Artes da Universidade Estadual de Feira de Santana – BA.<sup>47</sup>

#### 5.3.1 Projeto “Em busca das raízes do Português Brasileiro”

O projeto “Em busca das raízes do Português Brasileiro”, de natureza interinstitucional e interdepartamental, está inserido no grupo de pesquisas “Constituição, variação e mudança no/do português brasileiro da área da Linguística e subárea da Sociolinguística e Dialectologia pertencente ao Departamento de Letras e Artes da Universidade Estadual de Feira de Santana-BA, aprovado pelo CONSEPE UEFS (13/04/2009) e financiado pela UEFS/FAPESB Edital 001/2012, possui registro no CEP sob o Protocolo N° 140.511 (CAAE 04641412.7.0000.0053), e encontra-se em sua 3ª fase, com uma agenda de pesquisas voltadas para a inclusão de trabalhos relacionados aos aspectos morfossintáticos da língua portuguesa.

O projeto “Em busca das raízes do Português Brasileiro” tem como premissa a defesa da necessidade de realização de novas pesquisas para melhor entender o fenômeno da variação

---

<sup>47</sup> O referido projeto foi concebido e coordenado pela professora Dra. Eliana Sandra Pitombo Teixeira, entre os anos de 2009 e 2016 estando, desde 2017, atualmente, sob a coordenação da prof. Dra. Silvana Silva de Farias Araújo (DLA-UEFS).

e/ou mudança linguística e, umas de suas principais preocupações é ampliar a comparação de fenômenos brasileiros com a de outras variedades que não apenas a europeia, visto que já existe uma boa quantidade de pesquisas sobre essas duas realidades.

Assim, a ampliação que está na base do projeto alicerça-se no fato de que as semelhanças entre o PB e as variedades africanas do português não se mostram apenas em termos sócio-históricos, mas também em termos linguísticos. No caso da realidade sociolinguística de Luanda, vale destacar que muitos de seus habitantes são oriundos de diversas regiões do país que foram em busca de melhores condições de vida, mais oportunidades de trabalho e segurança. Sua população é também composta por falantes nativos de línguas diversas, com prevalência, daquelas pertencentes ao subgrupo banto. Deste modo, o projeto tem por objetivo geral consolidar a pesquisa linguística na UEFS, com destaque à história do PB, do Português de Luanda e ainda aponta que “os resultados advindos com esta investigação contribuirão com o debate acerca da importância de aspectos sócio-histórico-culturais na configuração atual do português brasileiro e, especialmente, no tocante à importância do contato do português com línguas africanas (Araújo, 2017, p. 5).

No resumo do projeto, pode-se perceber a preocupação em realizar pesquisas empíricas sobre o Português Luandense, ampliando o tema de estudo para outros aspectos da morfossintaxe:

[...] poderão ser investigados temas, como o do uso da colocação pronominal, dos artigos, da expressão verbal etc. Esses estudos que caracterizarão traços importantes da gramática do português luandense poderão ser comparados com outros já realizados com dados do português do Brasil (PB) e de Portugal (PE), trazendo ainda mais contribuições para o debate acerca da formação do PB (Araújo, 2017, p. 2).

### **5.3.2 As entrevistas**

As entrevistas seguiram um roteiro pré-elaborado, com base em informações sobre a cultura do lugar, mas sem a rigidez do padrão “pergunta/resposta”. A constituição da amostra foi feita a partir da seleção aleatória dos informantes, todos adultos e maiores de 21 anos de idade, seguindo, no entanto, alguns critérios que pudessem controlar melhor os fatores sócio-históricos. Para a seleção dos possíveis informantes, as pessoas entrevistadas deveriam ser oriundas de Luanda, ou que tenham, pelos menos, sido levadas para lá desde os primeiros anos de vida. Esse critério é bastante importante para a análise já que pode representar de forma mais fiel os aspectos e características linguísticas da comunidade pesquisada.

A coleta foi realizada nos anos 2008 e 2013 pela pesquisadora Eliana Sandra Pitombo Teixeira, *in loco*, na época, também coordenadora do Projeto “Em busca das raízes do Português Brasileiro” do NELP e contou com a colaboração de estudantes universitários e nativos de Luanda, como professores da comunidade local que eram responsáveis em fazer a intermediação entre os pesquisadores e os entrevistados. Esta estratégia de usar residentes foi pertinente pois pretendia-se que as pessoas entrevistadas se sentissem bastante à vontade com as perguntas que seriam feitas e que tivessem confiança para participar da pesquisa com uma pesquisadora principal até então desconhecida deles.

Com isso em vista, as entrevistas sociolinguísticas tipo DID foram conduzidas da forma mais natural possível para que não acarretasse nenhum tipo de constrangimento ou mal-estar ao participante pelo pesquisador responsável e pelos estudantes e professores universitários colaboradores de Luanda (sob a instrução e orientação da pesquisadora principal). Os inquéritos partiram de alguns questionamentos para que os participantes pudessem falar um pouco das suas culturas, sua escolaridade, suas línguas, de onde eram, seus costumes, o que faziam em relação ao trabalho e tempo livre e alguns acontecimentos marcantes, fatos mais delicados de se tratar. Diante disso, a preocupação em criar um ambiente agradável era crucial para os entrevistadores e entrevistados, pois, além de ajudar no andamento e fluidez da entrevista, no relato dos fatos, ao mesmo tempo, os entrevistados reproduziriam o vernáculo conforme os modelos labovianos, com uma fala menos monitorada e sem preocupações de aproximar a uma língua considerada mais culta.

Quando, geralmente, participantes são entrevistados, é natural que as pessoas se preocupem mais com as suas falas e tendem a aproximar de falas consideradas mais cultas. Para os participantes com níveis mais baixos de escolaridade, isso poderia ser algo ainda bem mais complicado visto que alguns deles poderiam se sentir inferiorizados e/ou incapazes de se comunicar com pessoas que possuem um grau de instrução escolar mais altos acarretando certa inibição durante a entrevista.

As entrevistas realizadas podem, de fato, ser consideradas conversas informais visto que não se tinha um questionário tão fixo de perguntas a seguir. Deste modo, à medida que a conversa ia avançando, outros tópicos, outros temas nas entrevistas iam aparecendo, evitando a formalidade, a rigidez das entrevistas baseadas apenas em perguntas e respostas. Segundo o projeto elaborado (Teixeira, 2007, p. 8), convém ainda destacar que o roteiro de entrevista, previamente elaborado, acabou sofrendo algumas adaptações conforme o desenrolar das entrevistas e às circunstâncias em que as entrevistas foram realizadas.

Para ajudar o trabalho dos inquiridores, foi necessário criar fichas com dados sociais dos informantes para um melhor controle da amostra com dados incluindo nome completo, idade, local de nascimento, referência da moradia etc. Caso fosse necessário, os pesquisadores recontatarem os participantes, ou apenas alguns deles, a ficha seria novamente consultada.

Com base nas conversas realizadas entre pesquisadores e pesquisados, pode-se encontrar respostas para as seguintes perguntas, separadas em blocos de interesses:

**Quadro 9: Blocos de Interesse (entrevistas)**

<b>Categorias</b>	<b>Perguntas</b>
<b>Cotidiano e comportamentos</b>	Qual é o veículo de comunicação que você mais utiliza para se manter informado? Você gosta de algum programa de TV? O que faz para se divertir em seu tempo livre? Onde costuma ir, o que gosta de comer e cozinhar e como prepara o alimento? Você é casada(o)? Tem filhos? Como é criá-los diante da sociedade? Como está acontecendo a educação das crianças nas escolas? Você acha que os filhos escutam os pais da mesma forma que vocês escutavam seus pais antigamente? Como os filhos se comportam quando os pais chamam a atenção?
<b>Infância</b>	Como foi a sua infância? Quais brincadeiras eram feitas? Quais músicas e histórias você costumava ouvir e se recorda? Quais acontecimentos mais marcaram a sua infância?
<b>Trabalho/ocupação</b>	Qual é a sua ocupação? Como é feito o seu trabalho diário? Você está satisfeita(o) com a profissão escolhida?
<b>Informações básicas</b>	Qual é o seu nome? Qual é o seu nível de escolarização? Qual é a sua língua materna? Qual é a sua idade? Qual é o seu local de nascimento? De onde são seus pais? Estudaram até que série? Por quanto tempo mora na capital Luanda? Você moraria em um lugar diferente?
<b>Línguas que são faladas</b>	Quando aprendeu a falar o português? Quais outras línguas africanas você fala? Em quais situações você utiliza a língua portuguesa e quando utiliza as línguas africanas? Com quem utilizam? Qual a importância do ensino de línguas nacionais na escola?

Fonte: elaboração própria

Um outro aspecto importante na condução das entrevistas foi o cuidado para que os participantes não se sentissem constrangidos diante de gravadores de áudio. Deste modo, foi

necessário que as entrevistas acontecessem nos mais variados locais como em escolas, nas casas dos informantes (no interior, quintais, varandas), nas casas de farinha e até em pastos e plantações. Muitas vezes, na tentativa de estabelecer um ambiente mais informal possível em que os entrevistados se sentissem bem em participar da pesquisa, em algumas gravações, foi possível ouvir a presença de pessoas externas e alguns outros ruídos não desejados como foi o caso do barulho da chuva, do choro de crianças, do barulho do vento no gravador etc. Mesmo com algumas interferências, o material teve uma boa qualidade das gravações realizadas.

A amostra das entrevistas do projeto aqui citado foi estratificada pela pesquisadora principal, época da coleta dos dados, com base em 5 variáveis sociais: sexo, faixa etária, nível de escolaridade, língua materna e local de nascimento dos informantes. Essa estratificação, ao longo desses últimos anos, tem servido de base para análises de diversos fenômenos linguísticos presentes na capital Luanda, conforme apresentado no quadro abaixo:

**Quadro 10: Estratificação da Amostra das entrevistas do Projeto Em busca das raízes do Português**

**Brasileiro**

SEXO	Feminino Masculino
FAIXA ETÁRIA	Faixa I (21 a 35) Faixa II (36 a 51) Faixa III (acima de 52 anos)
NÍVEL DE ESCOLARIDADE	Baixa ou Nula Médio Superior

Fonte: elaboração própria

O rico acervo de fala do projeto “Em busca das raízes do Português Brasileiro”, desde a sua idealização, tem contribuído para que novas investigações sobre a realidade sociolinguística de Luanda em seus diferentes campos, a exemplo do lexical, do semântico-pragmático, do morfossintático, entre outros. Espera-se que os objetivos propostos pelo Projeto continuem sendo alcançados e que possam servir de referência para outros estudos que busquem conhecer não só a variedade angolana da língua, assim como a brasileira, com suas similaridade e diferenças.

### 5.3.3 A transcrição

Após o momento da realização das entrevistas (DID), todo o material coletado foi, em sua maioria, transcrito grafematicamente pela equipe do Projeto já detalhado na subseção anterior. As transcrições realizadas foram feitas pelos pesquisadores e, principalmente, pelos bolsistas de Iniciação Científica (doravante IC) ou por membros do NELP/UEFS. As transcrições realizadas garantiram os registros das marcas linguísticas dos falantes, das expressões com marcas específicas dos seus falares, suas características fônicas e morfossintáticas. Buscou-se realizar as transcrições mais fiéis possível mesmo que algum dado não correspondesse à ortografia oficial. Nada foi desconsiderado.

### 5.3.4 A comunidade: Luanda

Luanda foi fundada em 25 de janeiro 1576 pelo primeiro governador de Angola, o explorador Paulo Dias de Novaes e só elevada à posição de cidade em 1605. Em 1648, após alguns anos do domínio holandês na região, Correia de Sá nomeou o local como “São Paulo da Assunção de Luanda”. Com a cidade que surgia, também iam aparecendo as primeiras fortificações e edificações.

O povoamento da cidade aconteceu muito em torno do tráfico de escravizados devido ao comércio bastante lucrativo e, por muito tempo, sofreu com grandes levadas de pessoas para regiões ultramar, longe de suas origens e família para suprir os desejos políticos e econômicos dos colonizadores nos diferentes tipos de lavoura como o café, cacau, cana-de-açúcar. Só no século XX, museus, instituições de ensino foram criadas, a exemplo do Museu de Angola (1938), do Instituto Angolano de Educação e Serviços Sociais (1962) e dos Estudos Gerais Universitários (1962).

Por ter uma localização privilegiada, no passado, foi e ainda é destaque portuário. Aos poucos, Luanda vem, cada vez mais, continuando a ter destaque, agora não mais como fonte de recepção e fornecimentos de negros para alimentar um comércio sub-humano, mas sim, trazendo de volta o crescimento e o progresso que o seu povo merece. Luanda é o maior centro urbano e econômico da república de Angola, localizada na região centro-norte do país além de estar situada no litoral oeste do Continente africano, na costa do Oceano Atlântico. Na Figura 15 é possível visualizar a capital nacional de Angola.

Figura 15: Localização cidade de Luanda (Angola) no território africano



Fonte: [https://www.researchgate.net/figure/Figura-1-Localizacao-de-Angola\\_fig1\\_299547757](https://www.researchgate.net/figure/Figura-1-Localizacao-de-Angola_fig1_299547757)

É após a Independência, depois de um período marcado por intensas guerras civis no interior do país, ocorrido entre os anos de 1975 e 2002, a cidade se transforma. Luanda passa então a receber um fluxo grande de pessoas que imigraram de diferentes regiões mais interioranas em busca de melhores condições de vida, de segurança e de saúde. As oportunidades que a capital poderia oferecer à população angolana no pós-guerra fez com que essa região fosse intensamente povoada por falantes nativos de diversas línguas, como já pontuado nesse trabalho, em sua maioria, pertinentes ao grupo banto, conforme pontuam Araujo e Dantas (2017, p. 70):

Além do contato entre línguas diversas, a história de contatos dialetais é uma tônica na cidade, pelo fato de ter recebido muitas pessoas de diferentes partes do país, em busca de segurança e de trabalho, principalmente a partir das lutas pela independência, nos anos de 1961 a 1975 e da Guerra Civil, que se estendeu de 1975 a 2002.

De acordo o Censo populacional de 2018, Luanda, capital angolana, tem cerca de quase 2,5 milhões de pessoas, sendo considerada uma das mais populosas do mundo. Hoje, a capital conta com a presença de muitas indústrias, desde o trabalho com produtos agrícolas, bebidas, alimentos processados até de fábricas que realizam montagem de carros etc, A cidade é considerada o principal centro financeiro, comercial e econômico do país. Há atividades de extração e refinaria de petróleo e é também considerada uma grande exportadora de produtos como o café, algodão, açúcar, sal e minérios, como diamantes e ferro.

Luanda também se destaca como polo universitário do país, abrigando universidades e institutos superiores de educação, não só públicas, mas também privadas. Dentre elas, está a Universidade Agostinho Neto (UAN), a maior e a mais antiga de todas. Outras instituições públicas da cidade são o Instituto Superior de Ciências da Educação de Luanda (voltado às licenciaturas); o Instituto Superior para as Tecnologias da Informação e Comunicação (voltada às engenharias e formações tecnológicas) e o Magistério Mutu-ya-Kevela (ensino técnico e formação de professores).

Mesmo as mulheres sendo um pouco mais da metade da população de Angola, em relação aos homens, muitas ações ainda precisam ser tomadas para que elas sejam mais empoderadas, não só socialmente, mas também, economicamente. Elas precisam ainda ter mais reconhecimento pelas atividades que exercem no mercado de trabalho em Angola. Sobre essa questão, Ana Paula Sacramento (2024), ministra da Ação Social, Família e Promoção da Mulher de Angola, em entrevista ao ONU News<sup>48</sup>, afirma que é de suma importância melhorar as condições de vida da população feminina, avançar na equidade de gênero para que as mulheres também tenham uma presença forte em diversos espaços da sociedade, desempenhando atividades mais formais, já que o mercado informal para elas ainda é predominante. Para ela:

“[...] As mulheres estão ainda no mercado informal, num bom número, mas o nosso trabalho continua no sentido de haver mais mulheres a saírem da linha da pobreza para terem uma condição de vida mais favorável” [...] elevar a condição da mulher com formação, informação, instrução, preparação, fazendo com que elas estejam também no centro da economia do país”.

A capital é ainda uma das 18 províncias de Angola. O fato de as províncias se subdividirem em partes ainda menores, faz com que Angola tenha um total de 162 municípios e 559 comunas<sup>49</sup>. As províncias de Luanda estão representadas na Figura 16, a seguir:

---

<sup>48</sup> <https://news.un.org/pt/story/2024/03/1829112>

<sup>49</sup> As comunas em Angola são o terceiro-nível de uma subdivisão administrativos e seguem os municípios em Angola.

**Figura 16: As 18 províncias de Angola, 2016**



Fonte: [https://pt.wikipedia.org/wiki/Prov%C3%ADncias\\_de\\_Angola](https://pt.wikipedia.org/wiki/Prov%C3%ADncias_de_Angola)

Desde modo, diante do que foi exposto até aqui, há de se considerar que Luanda, sendo um dos centros que mais se destaca no país em seus aspectos econômicos, políticos, educacionais e, inclusive populacionais, precisa ser investigada. Existe uma confluência entre diferentes línguas e culturas e, essa miscigenação pode nos ajudar a entender como essas pessoas, como suas histórias, como seus intensos contatos, podem influenciar a variedade portuguesa em Angola e que podem contribuir de alguma forma para os debates e entendimentos sobre as raízes do português brasileiro, algo que se pretende alcançar com os resultados dessa pesquisa, em especial, sobre o fenômeno de indeterminação aqui pesquisado visto que a fala como também a escrita refletem a heterogeneidade linguística da comunidade que habita a cidade e, quanto mais diversa, mais rico o estudo pode se tornar.

### **5.3.5 Os participantes da pesquisa**

No quadro 11, a seguir, são apresentados os perfis dos participantes da pesquisa no *corpus* utilizado. O quadro diz respeito aos falantes que aprenderam o Português como língua materna – L1. Uma parte dos colaboradores, ainda jovens, foi para Luanda por causa das guerras que assolavam o país e por busca de melhores condições de vida, de oportunidades de trabalhos, de estudos etc. Nas entrevistas, muitos deles informam que tem uma certa proficiência em alguma língua local (autóctone), mas muitos aprenderam a falar o português desde pequenos. Importante chamar a atenção para o fato de que muitos deles aprenderam uma língua africana depois de terem aprendido o próprio português.

No quadro também é mostrado a escolaridade em que cada um dos participantes se encontra, o sexo e a faixa etária.

**Quadro 11: Participantes Corpus Luanda (Português L1)**

Participante	Português L1		
	Escolaridade	Sexo	Faixa / Idade
1. P01.M1B	Nula e Baixa escolaridade	Masculino	I
2. P02.M1B	Nula e Baixa escolaridade	Masculino	I
3. P03.F1B	Nula e Baixa escolaridade	Feminino	I
4. P04.F1B	Nula e Baixa escolaridade	Feminino	I
5. P05.M2B	Nula e Baixa escolaridade	Masculino	II
6. P06.F2B	Nula e Baixa escolaridade	Feminino	II
7. P07.F2B	Nula e Baixa escolaridade	Feminino	II
8. P08.M3B	Nula e Baixa escolaridade	Masculino	III
9. P09.F3B	Nula e Baixa escolaridade	Feminino	III
10. P10.F3B	Nula e Baixa escolaridade	Feminino	III
11. P11.M1M	Ensino Médio	Masculino	I
12. P12.M1M	Ensino Médio	Masculino	I
13. P13.F1M	Ensino Médio	Feminino	I
14. P14.F1M	Ensino Médio	Feminino	I
15. P15.M2M	Ensino Médio	Masculino	II
16. P16.M2M	Ensino Médio	Masculino	II
17. P17.F2M	Ensino Médio	Feminino	II
18. P18.F2M	Ensino Médio	Feminino	II
19. P19.M3M	Ensino Médio	Masculino	III
20. P20.M3M	Ensino Médio	Masculino	III
21. P21.F3M	Ensino Médio	Feminino	III
22. P22.M1S	Ensino Superior	Masculino	I
23. P23.M1S	Ensino Superior	Masculino	I
24. P24.F1S	Ensino Superior	Feminino	I
25. P25.F1S	Ensino Superior	Feminino	I
26. P26.M2S	Ensino Superior	Masculino	II
27. P27.M2S	Ensino Superior	Masculino	II
28. P28.F2S	Ensino Superior	Feminino	II

29. P29.F2S	Ensino Superior	Feminino	II
30. P30.M3S	Ensino Superior	Masculino	III
31. P31.F3S	Ensino Superior	Feminino	III
32. P32.F3S	Ensino Superior	Feminino	III

Fonte: elaboração própria

Vale ressaltar que o projeto “Em busca das raízes do Português Brasileiro”, é composto por dois *corpora*, o de L1 e de L2<sup>50</sup>, com 59 entrevistas de participantes no total, no entanto, para este estudo, optou-se por analisar a língua portuguesa daqueles que primeiro aprenderam essa língua, desse modo, utilizou-se os dados dos falantes que nasceram em províncias próximas e que logo foram para a capital ou que sempre viveram em Luanda, visando analisar o vernáculo dos falantes em Luanda para que fosse possível obter uma fotografia, um panorama da situação linguística da comunidade em relação ao fenômeno da indeterminação do sujeito. Para futuros estudos ligados ao projeto destacado, convém ampliar a análise também para os falantes de L2, comparando os resultados desses dois *corpora*. Sendo assim, a L1 e L2 devem ser consideradas nos estudos de Luanda, bem como a localidade de origem (atrelada à língua da etnia de origem), as demais províncias, áreas mais afastadas do litoral já que a cidade recebeu milhares de pessoas de outras províncias.

A amostra composta pelos falantes de Português como L1, totalizou 32 entrevistados (15 homens e 17 mulheres), sendo que 10 deles nunca frequentaram à escola ou tinham pouca escolaridade, 11 pertencentes ao Ensino Médio e 11 do Ensino Superior. Em relação à faixa etária, 22 entrevistados pertencem aos dois primeiros níveis, a faixa I - adultos de 21 a 32 anos e a faixa II – adultos com idades intermediárias, de 33 a 52 anos, com 11 falantes cada e 10 falantes na faixa III, faixa etária que inclui os mais velhos, com 52 anos ou mais.

Na sequência, os dados aqui apresentados são mais bem visualizados no Quadro 12:

**Quadro 12: Visão resumida dos Informantes (Português L1)**

	<b>PORTUGUÊS L1</b>
<b>Nula/Baixa Escolaridade</b>	10
<b>Ensino Médio</b>	11
<b>Ensino Superior</b>	11
<b>Sexo</b>	<b>M- 15 F- 17</b>

<sup>50</sup> O *corpus* composto pelos falantes de Português como L2 é composto por entrevistas de 27 participantes (15 homens e 12 mulheres). Em relação à escolaridade, há 10 pessoas que tem nula ou baixa escolaridade, 09 que cursaram o Ensino Médio e 08 deles, o Ensino Superior. Ainda sobre o total de participantes, 10 deles se encontram na faixa etária I, 12 na faixa etária II e 05 na faixa etária III.

<b>Faixa etária I - (20 a 35 anos)</b>	11
<b>Faixa etária II - (36 a 51 anos)</b>	11
<b>Faixa etária III - (52 + anos)</b>	10
<b>TOTAL:</b>	32

Fonte: elaboração própria

### 5.3.6 Seleção das variáveis

É interessante pontuar que a escolha para a realização de algumas formas linguísticas pelos falantes não ocorre de forma aleatória na concorrência das variantes, os contextos em que elas acontecem são determinados. Diante disso, na constituição do envelope de variação desta pesquisa, foram considerados como base estudos prévios realizados sobre as estratégias de indeterminação do sujeito em amostras de fala e escrita, principalmente, realizados em território brasileiro, e poucos sobre a variedade da África (Angola) devido à sua escassez, algumas variáveis: as dependentes, as linguísticas e as sócio-históricas.

Sobre o subitem (iii), que trata das variáveis elencadas para este estudo, a variável dependente, no quadro 13, são apresentados alguns exemplos<sup>51</sup> encontrados em pesquisas realizadas anteriormente em relação ao fenômeno aqui investigado:

**Quadro 13: Variável dependente e seus respectivos fatores**

<b>∅ + V3PS</b>	Verbo 3ª pessoa do singular	∅ coloca duas juntas di boi né? (Pereira, 2013)
<b>∅ + V3PP</b>	Verbo 3ª pessoa do plural	Sim, já <b>dispararam</b> contra mim só que prontos, o fator sorte permaneceu. (Teixeira e Almeida, 2011)
<b>∅ + V+SE</b>	Verbo 3ª pessoa do singular + SE	[...] <b>nota-se</b> um crescimento muito grande da mulher [...](Teixeira e Almeida, 2011)
<b>∅ + VINF</b>	Verbo infinitivo impessoal	Tem qui fazê de tudo [...] ∅ lavá ∅ passa... ∅ cozinhá, ∅ arrumá... tudo isso (Pereira, 2013)

<sup>51</sup> A variável dependente encontra-se no eixo pragmático discursivo e necessitam de um maior contexto da entrevista, por isso, optou-se a não inserir todo o contexto, apenas fragmentos, já que a referência, muitas vezes pode estar mais distante do trecho em questão, alongando demais esse trabalho.

<b>VPSA</b>	Voz passiva sem agente	“Como nos <b>tem sido pedidas</b> algumas as-   signaturas do jornal para diversos pontos da   Europa [...] (Dias, 2017) mudar exemplo
<b>Nós</b>	Pronome 1ª pessoa do plural	Anteriormente <b>nós</b> criamos e metemos na escola. (Teixeira e Almeida, 2011)
<b>Você</b>	Pronome na 2ª pessoa do singular (TU)	“Entra para cá fulano.” Chegava lá <b>você</b> bebia, bebia [...] (Assunção, 2012)
<b>A gente</b>	Pronome na 1ª pessoa do plural (NÓS)	No mês de mauço, que é o mês que a gen... <b>a gente</b> chama aqui mês da quaresma, né? (Santana, 2006)
<b>Eles</b>	Pronome na 3ª pessoa do plural	[...] na questão de São Paulo que <b>eles</b> não gostam de nordestino que diz que nor- que nordestino só vão[...] (Mendonça; Nascimento, 2015)
<b>Eu</b>	Pronome 1ª pessoa do singular	Eu acho assim: se <b>eu</b> chego em sua casa pa fazer um serviço, desde que você sai, tem gente trabaiano em sua casa, se <b>eu</b> estou trabaiano nela, [...] (Assunção, 2012)

Fonte: Elaborado pela autora com base em exemplos de pesquisas já realizadas.

Com base no quadro 13 apresentado, vale destacar que há variáveis controladas que não possuem sujeito lexicalmente preenchido como é o caso dos contextos de  $\emptyset + V3PS$ ,  $\emptyset + V3PP$ ,  $\emptyset + V+SE$ ,  $\emptyset + VINF$ , enquanto outros já possuem o preenchimento da posição de sujeito SN como verificados em: *você*, *a gente*, *eles*, *eu* e *nós*. Convém lembrar que, a indeterminação do sujeito também pode ser identificada partindo de formas nominais ou “sintagmas nominais”, com exemplos já encontrados em alguns estudos anteriores, conforme se vê no quadro 14, a seguir.

**Quadro 14: Variável dependente – formas nominais**

<b>O homem</b>	<b>A pessoa</b> fumava, colocava as cinzas. (Guimarães, 2018)
<b>O sujeito</b>	quando <b>o sujeito</b> estava fugindo e o motorneiro alucinado aí ele me tirou botou num táxi (Carvalho, 2011)
<b>O povo</b>	<b>O povo</b> chamava de caramelo. (Guimarães, 2018)
<b>O indivíduo</b>	“E <b>o indivíduo</b> tinha o direito mas aí pagava por fora o médico né” (Carvalho, 2011)
<b>A pessoa</b>	[...] $\emptyset$ cria ele, depois, <b>a pessoa</b> cria[...] (Souza, 2014)

Fonte: Elaborado pela autora com base em exemplos de pesquisas já realizadas

As pesquisas apontam também para a presença de formas nominais como *a turma*, *o caba* (cabra), *o camarada*, *nego*, entre outros. Essas formas já documentadas na literatura bem como outras encontradas nesta pesquisa foram consideradas. A discussão sobre os sintagmas nominais é feita em outra subseção partindo de uma análise mais semântica do que morfossintática visto que também são consideradas expressões que ocupam posição de sujeito e têm interpretação indefinida, sendo mais difícil identificar o seu referente.

As variáveis controladas, as independentes, linguísticas e extralinguísticas são apresentadas no Quadro 15, a seguir, com alguns dos exemplos de casos de indeterminação do sujeito encontrados em pesquisas anteriores.

Quadro 15: Variáveis controladas

Grupos de Fatores	Fatores	Exemplos de indeterminação do sujeito
<b>G2: Grau de Indeterminação</b> <sup>52</sup>	<p><b>Indeterminação parcial com referência explícita no contexto</b></p> <p><b>Indeterminação parcial com referência implícita no contexto</b></p> <p><b>Indeterminação completa</b></p>	<p>Mai no tempo que a gente era rapaz, que a gente juntava o grupo [...]. (Santana, 2006)</p> <p>Já. Eu já e muito. Quano <b>a gente</b> vai assim tirar casca, a gente sobe num morro arto. (Santana, 2006)</p> <p>Não senhora, nunca não. Só de... desse bichinho que <b>Ø chama</b> la... lacraia. (Santana, 2006)</p>
<b>G3: Função da Indeterminação</b> <sup>53</sup>	<p><b>Desfocalização do sujeito</b></p> <p><b>Exemplificação</b></p> <p><b>Descomprometimento</b></p> <p><b>Ocultação do sujeito</b></p>	<p><b>Ø</b> chamá di mulequi eli num atendi. . . .(Pereira, 2013)</p> <p><b>Ø</b> posso isperá quando dá dá nos fim das colheita. . . (Pereira, 2013)</p> <p><b>nóis</b> dexava us mininu trancadu pra mim í trabalhá pur dia porque num podia trabalhá pur mês fichadu né?. . . (b) <b>Ø dexá</b> criança piquena era tudu piquininhu u mais novu tinha u quê? (Pereira, 2013)</p> <p>[...] “a responsabilidade não ia acompanhada de   folha corrida, quando aliás não ha quem   ignore que taes</p>

<sup>52</sup> Como será descrito mais adiante, o grau de indeterminação aqui referenciado leva em consideração ao proposto por Milanez (1982), já que a indeterminação, segundo a autora, não é de natureza absoluta e comporta graus quanto à sua abrangência. Essa análise do grau de indeterminação é largamente utilizada também por outros pesquisadores como Cunha (1993), Santana (2006), Dias (2017), entre outros.

<sup>53</sup> A função da indeterminação utilizada nesse estudo foi descrita por Milanez (1982) e complementado por Pereira (2013) ao incluir a função da “economia linguística”. Os autores têm utilizado a variável função da indeterminação no intuito de estabelecer as formas utilizadas a algumas das funções exercidas pela indeterminação e servindo como hipóteses explicativas para usos das estratégias de indeterminação.

	<b>Economia Linguística</b>	formalidades se dispensam   quando <b>se escreve</b> em causa própria.   Foi isto no dia 15 de junho.” (Dias, 2017)  “num tem mais matinê. . . (a) <b>nóis</b> tinha um cine/ cinema aí na brasilândia num tem mais” (Pereira, 2013)
<b>G4: Tipo de frase</b>	<b>Afirmativa</b>  <b>Negativa</b>  <b>Interrogativa</b>	<b>Vendeu</b> laranja, ‘bacaxi, no máquina de ferro [...] (Ponte, 2008)  <b>A gente num</b> se arrisca não. (Ponte, 2008)  [...] e hoje aqui já <b>ø faz</b> não é? (Assunção, 2012)
<b>G5: Tipo de Oração</b>	<b>Oração absoluta</b>  <b>Oração coordenada</b>  <b>Oração principal</b>  <b>Oração subordinada</b>	<b>Você</b> tem a festa do mundo todo na cidade. (Santana, 2006)  <b>Tinha uma igreja véia, derrubaro, fizeram ôtra</b> [...] (Santana, 2006)  <b>A gente pensa</b> que a gente tá passano na mesma rua (Santana, 2006)  Tão matano <b>como mataro um fio meu</b> (Santana, 2006)
<b>G6: Formas nominais do verbo</b>	<b>Finito</b>  <b>Infinitivo</b>  <b>Gerúndio</b>	[...] tudo <b>a gente</b> plantava. (Assunção, 2012)  “Tudo que <b>Ø prantar</b> aqui dá.” (Santana, 2006)  “[...] violam-se as consciências, <b>violentando-se</b> os mais puros intuítos; (Dias, 2017)
<b>G7: Tempo e modo verbal</b>	<b>Presente do indicativo</b>  <b>Presente do subjuntivo</b>  <b>Pretérito perfeito do indicativo</b>  <b>Pretérito imperfeito do indicativo</b>  <b>Futuro do subjuntivo</b>	[...] você não <b>tem</b> os ônibus suficientes pra deslocar, né? (Setti, 1997)  [...] se eu chego em sua casa pa fazer um serviço, desde que <b>você saia</b> , [...] (Assunção, 2012)  [...] hoje em dia <b>você soletrou</b> , ø ta sabeno ler [...] (Assunção, 2012)  A gente <b>comprava</b> tudo em grande quantidade (Setti, 1997)  A gente num pode fazer nada disso c’um ninguém, pode não. Se <b>a gente tiver</b> uma festinha dentro de casa, uma brincadeirazinha[...]

	<p><b>Pretérito imperfeito do subjuntivo</b></p> <p><b>Futuro do pretérito</b></p>	<p>(Souza, 2014)</p> <p>[...] graças a Deus, quer dizer, que <b>nós botasse</b> nossa bomba, certo?[...] (Souza, 2014)</p> <p>porque ali mesmo na, na região onde eu moro a zona de arrombamento é muito grande, antes a gente <b>poderia</b> até chegar em casa mais tarde (Assunção, 2012)</p>
<b>G8: Transitividade verbal</b>	<p><b>Verbo copulativo</b></p> <p><b>Verbo transitivo</b></p> <p><b>Verbo intransitivo</b></p>	<p>A gente <b>fica</b> mais experiente [...] (Souza, 2014)</p> <p>Porque a gente <b>teno um estudo...</b> tem tantos trabalho[...] (Souza, 2014)</p> <p>Mas a gente sempre <b>trabaia</b> mais, entendeu? (Souza, 2014)</p>
<b>G9: Tipo de verbo</b>	<p><b>Dicendi</b></p> <p><b>Epistêmico</b></p> <p><b>Estado</b></p> <p><b>Ação</b></p>	<p><b>ø falaram</b> que iam fazer (Assunção, 2012)</p> <p>[...] <b>you</b> não tem muito que <b>pensá</b> grande [...] (Franceschini, 2011)</p> <p>O povo <b>era</b> mais pegado com Deus, entendeu? (Souza, 2014)</p> <p>Não tem cabimento <b>Ø dormir</b> na fila pra <b>Ø marcar</b> médico[...] (Souza, 2014)</p>
<b>G10: Estrutura do núcleo do predicado</b>	<p><b>Estrutura simples</b></p> <p><b>Estrutura complexa</b></p>	<p>[...] tudo a gente <b>plantava</b>. (Assunção, 2012)</p> <p>O que tá difícil é quando o dinheiro não chega, que só <b>podemos entregar</b> na escola (Teixeira e Almeida, 2011)</p>
<b>G11: Ausência e Presença de Preposição (SE)</b>	<p><b>Presença</b></p> <p><b>Ausência</b></p>	<p>[...] nem é preciso pau <b>pra se</b> inducá fias. (Dias, 2017)</p> <p>...] Tem cabimento <b>se</b> passar por dono das casas alheias? [...] (Dias, 2017)</p>
<b>G12: Concordância com o argumento interno do verbo (SE)</b>	<p><b>Concordância</b></p> <p><b>Não-concordância</b></p>	<p>[...] com mais sangue frio, <b>multiplicaram-se as utopsias   cadavéricas</b> [...] (Dias, 2017)</p> <p>em cujo seio <b>se   conta as centenas de membros</b>. (Dias, 2017)</p>
<b>G13: Posição do argumento interno do verbo (SE)</b>	<p><b>SN posposto</b></p> <p><b>SN anteposto</b></p>	<p>[...] <b>Multiplicaram-se as utopsias   cadavéricas</b> [...] (Dias, 2017)</p> <p>[...] <b>grandes fortunas se   fizeram</b>. (Dias, 2017)</p>

G14: flexão do verbo (SE)	<p><b>Tempo finito</b></p> <p><b>Infinitivo</b></p>	<p>“É o que <b>se vê</b> descrito em um sem número de trabalhos econômicos.” (Carvalho, 2010)</p> <p>“sem <b>se chegar</b> a conclusões peremptórias.” (Carvalho, 2010)</p>
G15: Preenchimento do sujeito	<p><b>Sujeito pronominal realizado (suj. pleno)</b></p> <p><b>Sujeito pronominal não realizado (suj. nulo)</b></p>	<p><b>Eles</b> chamam de beberão de água. (Guimarães, 2018)</p> <p><b>Ø</b> Chama caolho. (Guimarães, 2018)</p>
G16: Mudança/manutenção do referente/ Paralelismo	<p><b>Referente igual à estrutura anterior na série discursiva</b></p> <p><b>Referente diferente da estrutura imediatamente anterior na série discursiva</b></p> <p><b>1ª forma de uma série</b></p>	<p>[...] depois <b>a gente</b> joga a semente e aí <b>a gente</b> espera assim[...] (Souza, 2014)</p> <p>[...] <b>a gente</b> faz o beiju, <b>Ø</b> coloca a massa, <b>nós</b> tem [...](Souza, 2014)</p> <p>Só no remelo. Tem um veneno que <b>a gente</b> passa na roça e acaba, a gente passou cedo aí, e ainda vai ter um pouco ainda (Souza, 2014)</p>
G17: Forma antecedente / Paralelismo	<p><b>Precedida de a gente explícito</b></p> <p><b>Precedida de a gente implícito</b></p> <p><b>Precedida de eles explícito</b></p> <p><b>Precedida de eles implícito</b></p> <p><b>Precedida de FN implícito</b></p> <p><b>Precedida de nós explícito</b></p> <p><b>Precedida de nós implícito</b></p> <p><b>Precedida de você explícito</b></p> <p><b>Precedida de você implícito</b></p> <p><b>Precedida de Ø+V+SE</b></p>	<p>a gente pisa o milho, bem pisado, <b>Ø limpa</b> ele (Santana, 2006)</p> <p>a gente pisa o milho, bem pisado, <b>Ø</b> limpa ele pele num ficar com pele, com pele. Aí agora, <b>Ø lava bem lavado</b> (Santana, 2006)</p> <p>qu'eles nem anda aqui. Só <b>Ø anda</b> aqui em quatro e quatro ano. (Santana, 2006)</p> <p>eles mexe mais lá é com frô. As horta dele, que <b>Ø</b> deve mexer em muita horta de... de verdura assim. Agora <b>Ø deixou</b> as verdura (Santana, 2006)</p> <p>A pessoa fumava, ..., machucava, <b>colocava</b> as cinzas. (Guimarães, 2018)</p> <p>Nós tê fé em Deus, que se <b>Ø num ter</b> fé em Deus (Santana, 2006)</p> <p>Premêro nós levava galinha, <b>Ø</b> matava duas galinha, <b>Ø fritava</b> (Santana, 2006)</p> <p>você pranta ele, com noventa dia <b>você tá colheno</b>. (Santana, 20006)</p> <p>Cê vai, pisa aquele piqui todim pra tirar aquela massa. Aí depois <b>Ø</b> lava de novo, agora <b>Ø pega</b> aquela massa (Santana, 2006)</p>

	<p><b>Precedida de Ø+V3PP</b></p> <p><b>Precedida de Ø+V3PS</b></p> <p><b>Precedida de Ø+VINF</b></p>	<p>se estudava, n'ê pra se estudar não, que, n'ê pra... pra <b>Ø aprender...</b> (Santana, 2006)</p> <p><b>Ø</b> picaro-lhe fogo, <b>Ø jogaro</b> pa debaixo do chão. (Santana, 2006)</p> <p><b>Diz</b> que é os dentes mais fortes que <b>a gente</b> tem, né. (Guimarães, 2018)</p> <p><b>Ø subir</b> e <b>Ø montar</b> em uma sela. (Guimarães, 2018)</p>
<b>G18: Inclusão do falante</b>	<p><b>Inclusão do falante</b></p> <p><b>Não inclusão do falante</b></p>	<p>[...] trabaiano na roça... <b>A gente</b> sempre usa o cabelo mais de fora [...] (Ponte, 2008)</p> <p>[...]poque eu tava sentada aqui e <b>Ø bateu</b> assim [...] (Ponte, 2008)</p>
<b>G19: Sexo</b>	<p><b>Feminino</b></p> <p><b>Masculino</b></p>	<p>_____</p> <p>_____</p>
<b>G20: Faixa etária</b>	<p><b>Faixa I</b></p> <p><b>Faixa II</b></p> <p><b>Faixa III</b></p>	<p>_____</p> <p>_____</p> <p>_____</p>
<b>G21: Escolaridade</b>	<p><b>Baixa ou nula escolaridade</b></p> <p><b>Ensino médio</b></p> <p><b>Ensino superior</b></p>	<p>_____</p> <p>_____</p> <p>_____</p>
<b>G22: Língua de Intercomunicação</b>	<p><b>Fala apenas português</b></p> <p><b>Fala português e compreende/fala língua bantu.</b></p>	<p>_____</p> <p>_____</p> <p>_____</p>

Fonte: Elaborado pela autora com base em exemplos de pesquisas já realizadas.

### 5.3.7 Critérios de exclusão na amostra

Antes de partir para a análise, na seção 6, das ocorrências das estratégias utilizadas para indeterminar o sujeito na fala angolana, convém destacar das amostras de falas analisadas os dados que foram excluídos considerando os seguintes contextos:

1. Falas reportadas de terceiros, pois não necessariamente configurariam falas reais dos entrevistados:

(1) “uma das vezes eu deixei assim a torneira aberta. Foi o primeiro dia que eu cheguei. quem trabalhava aqui era a filha. Chegou dela, ela disse assim: “ah, **começam** arranjar pessoas que não sabem fazer nada, vê lá como é que ela deixou a torneira aberta” (P04.F1B)<sup>54</sup>

(2) “Antigamente avó me dizia assim: “minha neta, nós nos tempo que éramos com, com, com a, com os nossos pais, **você** pa saber, pa escolher um, um, um ton marido que vai casar com ele, tens que saber vere, porque há homens que você lhe vê, é homem, mas chega à noite, não é homem.” (P10.F1B)

2. Fala do documentador considerada “gatilho”, no momento das entrevistas, as quais poderiam induzir o participante a repetir a mesma variante linguística utilizada:

(3) Inf.: Aí é só...

Doc: Controlar a área

Inf.: “Ø Controlar a área” (P08.M3B)

3. Expressões idiomáticas:

(4) “vamos ter que pôr a mão na massa, **como se costuma dizer.**” (P22.M1S)

(5) “**Não se fala mais nisso**” (P22.M1S)

4. Marcadores discursivos frequentemente utilizados pelos participantes:

(6) “já que **pretende saber**, de Luanda [...]” (P22.M1S)

(7) “Porque a condição da mulher em qualquer sociedade e uma condição, óbvia **quer dizer**, não há atividade separada entre o homem e da mulher [...]” (P30.M3S)

5. A presença de variáveis dependentes em situações em que não havia a realização de algum verbo, portanto, não se caracterizando em algum tipo de oração:

(8) “mas eu não quis seguir este curso porque no ensino médio **nós**... Eu estudei numa escola de formação de professore em que tínhamos também que dar aulas práticas. (P25.F1S)

(9) “então é fácil **você**... esse ano, por exemplo, tó a estudar de manhã, vou a faculdade de manhã e à tarde tó aqui a trabalhar. Da pra conciliar. (P24.F1S)

6. Falas dirigidas ao entrevistador pelo participante:

<sup>54</sup> Lê-se: participante 04, sexo feminino, faixa etária 1 e escolaridade baixa.

Nos excertos apresentados, a identificação dos participantes é feita utilizando a seguinte ordem: **P** (participante); **Número** (código da entrevista); **F/M** (sexo – feminino e masculino); **1/2/3** – (faixa etária 1, 2 ou 3) e **B/M/S** (nível de escolaridade – baixo, médio e superior).

(10) “ $\emptyset$  **Interpretas** bem ou mal aquilo que eu vou lhe dizer, mas é procurar chamar a atenção para que a próxima vez não volte a acontecer porque nós tamos aqui nessa vida, tamos sujeito a erro., tamos susceptíveis a erros.” (P11.M1M)

(11) “**Você** imagina que eu tenho um filho a estudar lá no lado do Jumbo, todos os dias eu gasto quatrocentos, fim do mês eu gasto dez dinheiro e eu não ganho isso, tá a ver?” (P18.F2M)

#### 7. Formas indefinidas:

(12) “É uma cidade que **ninguém** mexe no outro, **ninguém**  $\emptyset$  tira [com si] as coisas do outro assim à toa. (P15.M2M)

(13) “**Muitos**  $\emptyset$  optam pelo taxi, do que o transporte coletivo, por causa do atraso... (P5.M2B)

Finalizadas as descrições para a exclusão de alguns dados da amostra, descrevem-se as variáveis controladas na seção seguinte.

## 5.4 VARIÁVEIS CONTROLADAS

Nesta subseção, são apresentados os fatores da variável dependente, as independentes (linguísticas e extralinguísticas) que, com base em estudos prévios, foram selecionadas também para serem investigadas. Assim, buscou-se verificar as variáveis descritas, mas com a quantificação, a análise e a discussão dos dados (cf. seção 5), foi possível mostrar as variáveis mais favorecedoras e significativas para o estudo da indeterminação do sujeito por falantes luandenses.

### 5.4.1 Variável dependente

A variável dependente é composta por variantes e foi agrupada em dois grupos. As do primeiro grupo abarcam as formas encontradas nos compêndios gramaticais normativos que não têm o sujeito lexicalmente preenchido: o *verbo na terceira pessoa do singular acompanhado da partícula “se”* ( $\emptyset$ +V+SE), o *verbo na terceira pessoa do plural* ( $\emptyset$ +V3PP), o *verbo no infinitivo impessoal* ( $\emptyset$ +VINFIN), o *verbo na terceira pessoa do singular* ( $\emptyset$ +V3PS). Em todas elas, não fica evidente a presença de indicação de um agente a que possa ser referido. Há, assim, uma informação geral e indeterminada veiculada pela forma verbal, como, por exemplo, “ $\emptyset$  falaram bastante”. Sendo assim, para Milanez (1982), essas formas podem

expressar o mais alto grau de indeterminação, tendo assim uma indeterminação completa. Segundo a autora, isso acontece por causa do distanciamento de referências específicas a qualquer uma das três pessoas do discurso já que elas não são marcadas quanto à pessoa; e as que também possuem o sujeito preenchido lexicalmente como: *nós*, *a gente*, *você* e *eles*. Nesse contexto, as variáveis dependentes podem ter o preenchimento, com um núcleo não ocultado, a exemplo da frase “*Eles* comeram muitas maçãs”. Seguem exemplos dessas duas situações:

(14) “A pessoa ter pra contar né.” (P29.F2S) - Sujeito lexical preenchido

(15) “Ø indicaram a mim pra ser professora das mães”. (P21.F3M) - Sujeito lexical não preenchido

No segundo grupo, há as presenças das formas nominais (FN) / sintagmas nominais (SN), analisadas separadamente. São exemplos desse contexto:

(16) “Porque a luz divina indica... indica mesmo o caminho onde **a pessoa** deve andar”  
(P21.F3M)

(17) “**o pessoal** manda parar o taxi, antes desta passageira ou passageiro subir” (P21.F3M)

Nas subseções seguintes, são apresentadas cada variante apresentada nesta tese.

#### 5.4.1.1 Verbo na terceira pessoa do singular acompanhado da partícula “se” / Ø + V+SE

Sem fazer referência a pessoas determinadas, a terceira pessoa do singular acompanhada do pronome *se* é comumente utilizada pelas GT, como uma das formas clássicas para a indeterminação do sujeito, pois não faz referência a pessoas determinadas em construções com a presença de verbos intransitivos, transitivos diretos ou indiretos.

Conforme Bagno (2011, p. 813), devido à uma reanálise sintática pela qual o PB vem passando, o pronome *se*, cada vez mais, vem assumindo um papel de sujeito. Outro motivo que agrega esse novo contexto se deve ao fato de haver uma exigência de acompanhamento do sujeito ao seu verbo, afirma o autor. “Assim, na sintaxe brasileira, aumenta progressivamente a tendência a não se deixar nenhum verbo sozinho, desacompanhado de seu sujeito, mesmo quando esse verbo é um infinitivo tradicionalmente classificado de *impessoal*” (Bagno, 2011, p. 813, grifo do autor).

Dias (2017), por exemplo, verificou que, nos textos escritos baianos dos séculos XIX e XX, a forma sem o sujeito lexical preenchido acompanhado do *se* foi a estratégia mais frequente (quase 50% dos registros) em todas as funções de indeterminação, principalmente, nos casos

em que a indeterminação foi completa. “Tipicamente genéricas, isto é, de **sujeito** maximamente **indeterminado**, já que todas as **pessoas** do discurso ficam abrangidas, são as construções de **terceira pessoa do singular** com o **pronome SE** [...]” (Neves, 2000, p. 463- 465, grifos da autora). Essa situação descrita por Neves (2000) foi verificada nos resultados de Dias (2017), pois o  $\emptyset+V+SE$  foi a estratégia mais indeterminadora de um *continuum*, pois, quando se utilizou essa construção, a identificação do referente foi dificultada.

O *continuum* ao qual Dias (2017) retrata em sua tese faz referência ao contínuo de gêneros textuais já descritos por Marcuschi (2008). Dias, assim, ao analisar os gêneros epistolares (redigidos para serem lidos) distinguindo-se das peças teatrais (redigidas para serem faladas), considerou o tipo de gênero para a indeterminação do sujeito por acreditar que o escritor seleciona alguns tipos de estratégias a depender da finalidade do texto.

Tomando-se como base a variante *3ª pessoa do singular + se*, busca-se verificar em que medida essa forma, largamente aceita nos compêndios gramaticais, é encontrada na fala dos luandenses. Foi exemplo encontrado no *corpus* para essa pesquisa:

(18) “depois disso, **abriu-se** mesmo alfabetização” (P21.F3M)

#### 5.4.1.2 Verbo na terceira pessoa do plural / $\emptyset + V3PP$

O verbo na 3ª pessoa do plural sem seu sintagma nominal é uma das formas mais clássicas e reconhecidas pelas GTs para indeterminar o sujeito. Sobre o uso dessa variável, o estudo de Setti (1997), nas três capitais do Sul do Brasil, mostrou que a escolaridade foi relevante ao apontar que os falantes com mais tempo de escolaridade estão utilizando mais a forma consagrada dos que os falantes menos escolarizados. Esses últimos utilizaram mais formas com o sujeito preenchido, a exemplo de *ele*.

Neves (2011, p. 464) afirma que a forma pronominal mais usada e que possui uma referência genérica é o *eles* –  $\emptyset+V3PP$ . A autora ainda afirma que esse contexto pode ainda ser mais comum quando a referência genérica se faz com a terceira pessoa do plural sem que exista o uso do pronome sujeito, como em: “Jogaram alguém na piscina [...]” ou “[...] na qual sem dúvida pretendiam extorquir-me[...]”. Menon (1994) aponta as formas lexicalmente preenchidas (*eles*) e a categoria vazia ( $\emptyset+V3PP$ ) que são covariantes e, desse modo, indicam uma única forma de indeterminação do sujeito. Diferentemente, sobre os usos das estratégias *eles* e  $\emptyset+V3PP$ , Milanez (1982) afirma que não são concorrentes entre si, visto que a ausência do preenchimento pronominal *eles* não pode ser considerada opcional. Conforme a autora, a forma preenchida é utilizada pelos falantes quando há uma referência anterior, implícita ou

explícita a um grupo que o sujeito faz parte, enquanto a estrutura ØV3PP não tem esse tipo de restrição.

Como o Ø+V3PP é uma abordagem consagrada pelas GT, pretende-se descobrir até que ponto essa variante é utilizada nas escolhas realizadas pelos falantes em Luanda quando utilizam uma construção em que seu referente não é determinado. É considerado nessa pesquisa que essas duas formas *eles* e Ø+V3PP não são variantes assumindo, assim, papéis diferentes na indeterminação do sujeito.

#### 5.4.1.3 Verbo na terceira pessoa do singular / Ø+V3PS

Quando se trata das diversas estratégias para indeterminar o sujeito, a construção Verbo na terceira pessoa do singular acompanhado da partícula “se” (Ø +V+SE) não é uma das construções mais encontradas nas GTs.

Menon (1994, p. 254) declara que ainda é “muito instável” a fronteira que existe entre o infinitivo impessoal e o pessoal porque “a 3.a p.s. do infinitivo pessoal é morfologicamente não-marcada e por este fato idêntica à forma do infinitivo impessoal”. No entanto, a autora acrescenta que “a equivalência das duas formas reforça a interpretação do infinitivo como forma de indeterminação, o infinitivo impessoal transformando-se em pessoal e constituindo, por conseqüência, a forma ØV3PS.”

Neves (2011, p. 464) afirma que “menos comum e de registro mais popular é o emprego da **terceira pessoa do singular** para a **indeterminação do sujeito**” e exemplifica da seguinte forma “*Lá tira título de eleitor, documento* (HO)”. (grifos da autora). Os resultados do verbo na terceira pessoa do singular, desacompanhado da partícula ‘se’ ocorreu em 7,1% dos dados de Carvalho (2010) e 5% de Assunção (2012), o que confirma a afirmação de Neves (2011) sobre um uso não alargado para indeterminar. Vale destacar que Souza (2014), diferentemente dos resultados de Carvalho (2010) e Assunção (2012), obteve um percentual de 28% geral da sua amostra para essa variante, principalmente entre a população menos escolarizada.

Em Ponte (2008), a estratégia Ø + V3PS se apresentou bastante produtiva nas comunidades estudadas. Considerando que as entrevistas foram realizadas em comunidades rurais afro-brasileiras na Bahia, a autora, fundamentada pelos conceitos de TLI e da bipolarização em que se encontra realidade sociolinguística, confirma a hipótese de que os aspectos morfossintáticos dos dialetos rurais, ao mesmo tempo se distanciam e se aproximam da norma culta. Ela afirma que:

A aproximação entre o comportamento lingüístico desses dois pólos se dá pela dissolução das marcas mais fortes da aquisição imperfeita do PB na fala das camadas populares da sociedade e pelo afrouxamento dos padrões lingüísticos das camadas médias e altas, que se inicia com os processos de independência do país. Por causa do nivelamento lingüístico que atinge as variedades populares do PB, propiciado pela recente e intensa urbanização do país e pela vulgarização do sistema de ensino e da mídia, as características lingüísticas crioulizantes só são encontradas em poucas comunidades rurais afro-brasileiras isoladas. (SOUZA, 2014, p. 120)

É possível que, na comunidade luandense, devido às especificidades do processo colonizador, as estratégias preconizadas nas gramáticas normativas para a indeterminação do sujeito estejam mais presentes. Com a imposição da LP, a aquisição pelos falantes também aconteceu de forma imperfeita já que muitos angolanos tiveram que aprender o português já adultos e como segunda língua, sem falar nas interferências das suas línguas nativas, em sua maioria do grupo banto. Por essas, é interessante investigar a forma  $\emptyset + V3PS$  nessa tese. É exemplo da variante  $\emptyset + V3PS$  no *corpus*:

(19) “Mas só que não *encontrou* nada. Não ***encontrou*** nada em casa” (P21.F3M)

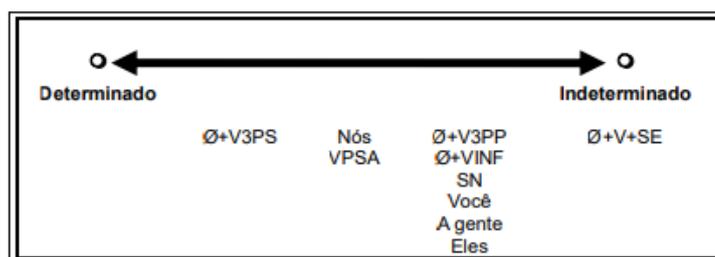
#### 5.4.1.4 Verbo no infinitivo impessoal / $\emptyset + VINF$

Quando um verbo se encontra no infinitivo impessoal ( $\emptyset + VINF$ ), não há como especificar o agente da oração, portanto, não se faz menção a qualquer uma das pessoas do discurso. Sendo assim, ele atribui um valor genérico ao sujeito como em (Santana, 2006), “Tudo que  ***$\emptyset$  prantar*** aqui dá”; “Tem que  ***$\emptyset$  trabalhar,  $\emptyset$  cozinhar***, cuidar da casa, né não?”, em uma cadeia discursiva que não é possível determiná-lo. Embora muitas pesquisas apontem uma frequência de uso pelos falantes para indeterminar o sujeito, muitas gramáticas, normalmente, não têm considerado essa estratégia como uma forma clássica para essa função. Dias (2017), ao analisar algumas gramáticas dos séculos XIX e XX, verificou que apenas algumas gramáticas faziam referência a esse tipo de variante. As três gramáticas que colocaram o infinitivo impessoal como uma possibilidade de indeterminar o sujeito foram a de Barbosa (1822; 1830), Rabello (1872) e Pereira (1907).

Nas gramáticas utilizadas como referência nesse estudo, percebeu-se que a estratégia  $\emptyset + VINF$  é retratada, nas gramáticas prescritivas de Luft (1996) e Cegalla (2008), e a descritiva de Perini (2002), como possibilidade de indeterminação do sujeito com o infinitivo.

Milanez (1982), por exemplo, reconhece essa variante como uma das formas que melhor expressa a generalização, já que é um tipo de construção despronominalizada. Desse modo, as formas,  $\emptyset$  3PS,  $\emptyset$ 3PS + se, além do  $\emptyset$  + VINF, “[...] distanciam-se de referências específicas a qualquer das três pessoas do discurso, uma vez que não são marcadas quanto à pessoa” (Milanez, 1982, p. 81). Em pesquisa sociolinguística, como a de Dias (2017), na intenção de representar um possível *continuum* para os graus de indeterminação, o autor criou um esquema ao considerar a frequência da distribuição das formas analisadas em sua tese da seguinte maneira:

Figura 17: *Continuum* (graus de indeterminação)



Fonte: Dias, 2017, p. 211

Conforme se vê,  $\emptyset$  + VINF, também têm contribuído de forma significativa na indeterminação do sujeito, muito embora, não sendo o maior favorecedor nos estudos já apresentados. Desse modo, analisou-se qual o seu grau de influência no *corpus* de Luanda. Segue, exemplo:

(20) “Não tinha professora ou como  $\emptyset$  dar atividade às senhoras e  $\emptyset$  passar atividade em si” (P21.F3M)

#### 5.4.1.5 Verbo na voz passiva sem agente (VPSA)

O verbo na voz passiva sem agente (VPSA<sup>55</sup>) é quando o agente não está expresso. É uma variante que, embora possa ser considerada pouco frequente como uma forma de indeterminação do sujeito, é ainda uma estrutura passível de ser achada. Conforme Haug (1986) afirma, a passiva analítica faz referência a qualquer pessoa do discurso, enquanto, a sintética referencia uma terceira pessoa.

<sup>55</sup> Não foram encontrados dados dessa variante na amostra.

Em Salvador, essa variante não foi encontrada por Carvalho (2010)<sup>56</sup> mas, em um estudo posterior, Dias (2017) encontrou o equivalente a apenas 0,7% dos recursos de indeterminação em seus *corpora*. Em um gráfico elaborado pelo autor referentes às estratégias de Indeterminação ao longo dos séculos XIX e XX, percebe-se que, em Dias (2017, p. 143), a passiva sem agente explícito vem decrescendo, vertiginosamente, não sendo considerada uma forma utilizada no material analisado.

Uma outra análise é descrita por Setti (1997, p. 62-63), ao afirmar que o pronome *se* com a função de indeterminação “(onde a voz passiva sintética pode se encaixar)”, o *VPSA* e o *ØV3PP* podem ser equivalentes. Para chegar a essa conclusão, a autora faz uma análise sobre as formas *se* e o *VPSA*, afirmando que o *se*, quando estiver acompanhado de um verbo que é transitivo, o sujeito é indeterminado como nos exemplos de Haury (1983, p. 171) utilizados por ela: "Resolveram-se os problemas. Desvendou-se o mistério.", visto que o sujeito da ação ou o seu responsável não se encontra explícito. Seguindo esse pensamento, partindo dos exemplos de Haury (1983), Setti afirma que, quando o *se* não for suficiente para indeterminar, é necessário realizar a conversão para a voz passiva analítica, como em: "Os problemas foram resolvidos. O mistério foi desvendado." Segundo a autora, ao fazer essa conversão, é possível ter um agente indeterminado e que, conseqüentemente, terá um sujeito indeterminado, na passagem para a voz ativa, a exemplo de “Resolveram os problemas. Desvendaram o mistério.”

#### 5.4.1.6 Formas com o sujeito lexical preenchido

Alguns estudos sobre as estratégias de indeterminação do sujeito apontam uma tendência ao uso de certas formas pronominais para indeterminar o sujeito pelos falantes, conforme Milanez (1982), Carvalho (2010), Assunção (2012). Sobre isso, Duarte (1993) aponta como uma mudança em progresso o uso de recursos pronominais expressos na indeterminação argumentando que a inserção do *a gente* no quadro dos pronomes pessoais, gerou uma redução no paradigma flexional dos verbos e, por consequência, o traço [+ pessoa] foi perdido no PB. Dessa forma, a manutenção do traço mais pessoal só aconteceria se os pronomes tivessem o sujeito explícito. A seguir, são explicitados os contextos linguísticos de *nós*, *você*, *a gente*, *eu*, *eles* para o sujeito lexical preenchido.

---

<sup>56</sup> O autor ressalta que “o que não quer dizer que ela não seja empregada pelos falantes de Salvador, uma vez que a base de dados constitui apenas uma amostragem.” Carvalho (2010, p. 77)

#### 5.4.1.6.1 Nós

Para as GTs, o *nós* é classificado como sendo o plural de *eu*. É o plural de mais outras pessoas, e não o eu + eu. Sobre isso, Benveniste (1995, p. 256), afirma que não existe a possibilidade do *eu* ser um processo de pluralização já que, na maioria das línguas o plural pronominal não equivale ao pronome nominal, assim, “se não pode haver vários ‘eus’ concebidos pelo próprio ‘eu’ que fala, é porque ‘nós’ não é uma multiplicação de objetos idênticos, mas uma junção entre o ‘eu’ e o ‘não-eu’, seja qual for o conteúdo desse ‘não-eu’”. Sobre isso, ele ainda acrescenta que “em ‘nós’ é sempre ‘eu’ que predomina, uma vez que só há ‘nós’ a partir de ‘eu’ e esse ‘eu’ sujeita o elemento ‘não-eu’ pela sua qualidade transcendente. A presença do ‘eu’ é constitutiva do ‘nós’” (Benveniste, 1995, p. 256).

Ainda conforme Benveniste (1995, p. 257), o *nós* não é um *eu* quantificado ou multiplicado, mas um eu “dilatado além da pessoa estrita, ao mesmo tempo acrescido e de contornos vagos.”. Desse modo, o *nós* é considerado um “eu ampliado” já que pode ser caracterizado por uma pessoa difusa e amplificada. Vale lembrar que, em algumas situações, o *nós* ainda pode ser utilizado para indicar um plural de modéstia quando se faz uma referência a si próprio sem deixar de fazer o mesmo também a qualquer outra pessoa com referência indeterminada e, evitando, assim, um modo impositivo ou arrogante nas afirmações e mostrando respeito. Segundo o autor, somente a terceira pessoa do discurso - ele -, caracterizada como uma não – pessoa (nem locutor, nem interlocutor, seres vivos ou objetos), aceitaria o plural verdadeiro já que ela não é marcada pela unicidade de pessoas do discurso. Neves (2000, p. 459) corrobora o fato de que os pronomes plurais considerados os de primeira pessoa, a exemplo do *nós*, sempre vão envolver um ‘não-eu’, sendo assim, extrapolam e não apenas fazem referência à primeira pessoa.

O pronome *nós*, como já comentado, brevemente, em outras seções, é apresentado como um recurso utilizado para indeterminar o sujeito no português brasileiro. Duarte (1995, 2000) já aponta esse contexto como uma forma de evidência de preenchimento da posição de sujeito por pronomes plenos, considerando assim uma língua de sujeito nulo parcial.

Na perspectiva do uso do recurso “nós” para indeterminar o sujeito, Santana (2006) exemplifica: “Tá, pois se *nós* tamo no fim da era. É de pegar, vai passano mermo aí, porque é pá nós ir embora tudo” e, podem também ser visto nos trabalhos de Milanez (1982), Menon (1994), Setti (1997), Godoy (1999), Dias (2017), Assunção (2012), Souza (2014), etc.

São exemplos da variável *nós* no *corpus*:

(21) “[..] **nós** víamos o tempo que as vezes [dois dias] não Ø conseguia sair fora. [...]”

(P29.F2S)

(22) “**Nós** viemos do centro da cidade, num é? (P10.F3B)

#### 5.4.1.6.2 A gente

Omena (2003) afirma que o substantivo *gente*, em *a gente*, passou a integrar o sistema de pronomes pessoais do português, concorrendo com o *nós* (P4) através do processo de gramaticalização. Há assim “uma mudança na gramática do item, que vai gradualmente cristalizando algumas das características do SN e perdendo outras, num processo de mudança de classe que se acentua com o acréscimo do traço semântico de primeira pessoa gramatical” (Omena, 2003, p. 64). A autora ainda mostra que o *a gente* conservou o traço generalizador, indefinido, que já era característico dele e passou a ser mais usado para indicar contextos de indeterminação e abrange um maior número de referentes até mais do que o *nós*.

O termo *a gente* também funciona como uma estratégia de indeterminação, pois não é possível identificar o referente ao qual a expressão remete, visto que a forma apresenta uma menor especificidade do sujeito. Em Ponte (2008), Assunção (2012) e Souza (2014), Guimarães (2018), a variante *a gente* mostrou-se mais produtiva, inclusive, se comparada às formas não pronominais. Santana (2014) e Souza e Oliveira (2014), com estudos realizados na Bahia, constataram que as mulheres apresentaram como tendência de estratégia de indeterminação o uso do *a gente*. Exemplo da variante *a gente* é:

(23) “A vida é memo assim né? À vez **a gente** vive um certo tempo [...]. (P10.F3B)

#### 5.4.1.6.3 Você

O pronome *você* resultou de um processo de gramaticalização da locução nominal Vossa Mercê e assumiu um *status* de pronome. Castilho (2019) afirma que essa gramaticalização do Vossa Mercê aconteceu em distintos campos, a saber: o fonológico (Vossa Mercê>vosmecê>você>ocê>cê; ou Vossa Mercê>vosmicê>vassuncê), o sintático (a reanálise do sintagma nominal como um pronome pessoal) e o pragmático: (antes tratamento dispensado aos reis, os novos-ricos quiseram o mesmo tratamento para eles, com a ascensão da burguesia)

Percebe-se que o pronome *você* apresenta uma desinência de 3ª pessoa do singular, mas, de fato, tem funcionado como 2ª pessoa e sua referência está mais ligada com quem se fala do que à pessoa de quem se fala. Na *Gramática de Usos do Português*, Neves (2011), afirma que

as formas *você* e *vocês* fazem referência à 2ª pessoa, no entanto, o verbo se encontra na 3ª pessoa. Segundo ela, o mesmo acontecia com os pronomes de tratamento, a exemplo de *vossa senhoria*, *vossa excelência*, *o(a) senhor(a)*. Para a autora, a forma *você* é largamente utilizada se comparada ao *tu*, principalmente na modalidade oral já que são usadas formas de segunda pessoa onde há o tratamento *você*, misturando as formas de 2ª e 3ª pessoas.

Enquanto a forma *você* é muitas vezes apontada pelas GTs como um pronome de tratamento que se refere à 2ª pessoa, Neves (2011, p. 463), em sua gramática descritiva, afirma não somente que os pronomes pessoais podem fazer referência genérica, como também inclui a expressão *você* nesse contexto. Para a autora, o pronome *você* pode indicar referência genérica e, nesse caso, a indeterminação é considerada muito forte pois o *você* pode indicar qualquer pessoa no discurso. Essa posição de Neves (2011) pode ser verificada nos estudos de Menon (1994), Setti (1997), Santana (2006), Assunção (2012), Souza (2014), entre muitos outros estudos que indicam o *você* é um recurso muito profícuo para indeterminar o sujeito.

Teixeira (2008, p. 7), pesquisando o *você* em Luanda, afirmou que, tanto na norma culta angolana quanto na portuguesa, “o pronome *tu* é usado como forma de tratamento íntimo, ficando o “*você*” para marcar distância ou, como afirmou um informante, é usado com pessoas nas quais o locutor não tem confiança ou simplesmente não quer dar confiança.” No entanto, ela mesma observou que o *você* estava sendo utilizado como um pronome de tratamento íntimo. Seus resultados mostraram que o *você* foi facilmente favorecido como pronome de segunda pessoa pelos falantes nativos de línguas nacionais, desde aqueles analfabetos ou com pouca escolaridade, até por falantes que aprenderam o português em situações mais adversas. A pesquisadora ainda compara o fato de o pronome *tu* no Brasil ser combinado com a desinência de 3ª pessoa e o mesmo tem acontecido com o *você* em Angola, pois, de alguma forma, a distinção entre 2ª e 3ª pessoas se perde.

A seguir, tem-se exemplo de *você* verificado no *corpus* investigado nesta tese:

(24) “**Você** pode falar o Português, mas nunca Ø vais considerar que fala bem o Português, [...]”. (P15.M2M)

#### 5.4.1.6.4 Eles

Milanez (1982) afirma que o pronome *eles*, ao fazer referência a um grupo, também pode assumir um valor genérico, pois não é possível identificar o seu referente, não possui nenhum referente no domínio discursivo presente dentro da estrutura linguística. Desse modo, a forma

pronominal *eles* pode acabar adquirindo a função de um sujeito que indetermina e que faz referência a um grupo mais amplo, podendo ser atribuído a qualquer pessoa dentro de um grupo.

Ainda, para a autora, as formas  $\emptyset+V3PP$  e *eles* são aquelas que possuem um menor grau de indeterminação já que são recursos exclusivos de terceira pessoa e, geralmente, são antecedidos por uma referência a um grupo social a que pertence o agente verbal. Para Menon (1994), a forma *eles* seria a forma plena da tradicional  $\emptyset+V3PP$  quando o uso é indeterminado, assim, seriam concorrentes considerando uma única forma. Essa visão difere do ponto de vista de Milanez (1982) e Godoy (1999) e Santana (2006), pois Menon acredita que não há distinção do uso entre essas duas formas. Outra justificativa dada pela autora é que o uso da forma pronominal *eles* acontece em decorrência de o fato da língua portuguesa estar preenchendo a posição de sujeito. Um exemplo de *eles* encontrado foi:

(25) “[...] então **eles** vão domingo e só voltam na sexta-feira. [...]” (P15.M2M)

#### 5.4.1.6.5 Eu

Menon (1994) destaca que o pronome *eu*, quando se torna indeterminado, perde o seu caráter dêitico e assume um valor equivalente ao *se*. Como uma das ilustrações para o *eu* genérico, a autora exemplifica esse contexto com uma fala de uma informante professora que relata sobre o período pré-histórico: “ora a maneira do homem pré-histórico era... Basicamente **eu** preciso comer... e **eu** preciso:: ... me defender dos animais e **eu** preciso me esquentar na medida do possível.” (Menon, 1994, p.204, grifos da autora). Nesse exemplo, só o distanciamento do tempo já fica evidente que o referente não pode ser a informante.

Neves (2011, p. 463-464) em “**EU** vou lá, fico dois dias fazendo curso, eles **ME** catequizam, **ME** fazem comprar uma tonelada de sabão e abrir o meu negócio.”, afirma que a princípio, o *Eu*, a forma característica de primeira pessoa que seria altamente determinada, acaba tendo também uma referência genérica, já que o falante pode imaginar qualquer pessoa a realizar a ação. Carvalho (2010), Assunção (2012), entre outros pesquisadores, analisaram o pronome *eu* no sentido generalizado, podendo ser referenciado a qualquer pessoa. Nesta pesquisa, essa variável também foi considerada para observar o seu caráter indeterminador na comunidade de Luanda.

#### 5.4.1.6.6 Formas nominais

As formas nominais são consideradas aqui “sintagmas nominais” (SN). Embora essas formas possam apresentar um sujeito expreso, ainda assim podem apresentar uma ideia mais genérica, mais generalizadora sobre o agente da ação. Os sintagmas nominais são expressões em que seus núcleos podem ser constituídos por nomes e ser antecidos por um determinante, como por exemplo, *as pessoas*, ou sem a presença do determinante. Conforme Bagno (2011), esses sintagmas nominais genéricos, termos e expressões cristalizadas socialmente se referem a ninguém em particular, mas sim, às pessoas em geral e, essas variam seus usos de acordo o grau de monitoramento. São elencadas as seguintes formas pelo autor Bagno (2011, p. 821): a criatura, a pessoa, beltrano, fulano (de tal), neguinho, o camarada, o cara, o cidadão, o cristão, o maluco, o outro, o pessoal, o povo, o sujeito, sicrano, o zé, zé dos anzóis carapuça etc. Muitas das formas analisadas por Bagno (2011) foram também encontradas em muitas pesquisas sociolinguísticas resenhadas na Seção 3 desta tese. De acordo Bagno (2011), ao analisar dados do projeto NURC-Brasil, por exemplo, as expressões **neguinho**, **o cara**, **o maluco** estavam ligadas a situações de extrema informalidade, sendo que a última forma (o maluco) foi utilizada em faixas etárias mais específicas. Algumas das formas nominais presentes no *corpus* analisado são:

(26) “**O pessoal** manda parar o taxi” (P21.F3M)

(27) “Ih, antes mesmo já... graças a Deus **a pessoa** sai de casa” (P21.F3M)

(28) “Os pais não... quando vê que **a criança** está indo mal os pais não dizem nada”  
(P21.F3M)

#### 5.4.2 Variáveis Independentes

A língua reflete comportamentos de uma comunidade e, sendo a língua um produto social, ela não pode ser separada das funções sócio-interacionais. A diversidade é uma propriedade funcional e inerente dos sistemas linguísticos e o estudo realizado pelo sociolinguista, segundo Labov (2008 [1972]), deve ser feito observando-se o seu contexto linguístico e extralinguístico.

Com o intuito de observar as variáveis sociolinguísticas que estão ligadas a esta pesquisa, as próximas subseções discorrem sobre as variáveis linguísticas e extralinguísticas relevantes para o estudo das formas de indeterminação do sujeito em Luanda com base em pesquisas prévias e do *corpus* utilizado.

#### 5.4.2.1 Variáveis Linguísticas

Nesta subseção são melhor explicadas cada uma das variáveis linguísticas selecionadas para análise, como: grau de indeterminação, função da Indeterminação, tipo de frase, tipo de oração, flexão do verbo (SE), tempo e modo verbal, transitividade verbal, tipo de verbo, estrutura do núcleo do predicado, ausência e presença de preposição, concordância com o argumento interno do verbo (SE), posição do argumento interno do verbo, preenchimento do sujeito, mudança/manutenção do referente, forma antecedente/paralelismo, inclusão do falante.

##### 5.4.2.1.1 Grau de Indeterminação do contexto

A análise do grau de indeterminação parte do que foi proposto por Milanez (1982). Quanto a essa variável, ela afirma que a “indeterminação não se apresenta como um fenômeno de natureza absoluta, que se opõe à determinação sempre de forma nítida; ao contrário, a mesma comporta graus quanto à sua abrangência” (Milanez, 1982, p. 80, grifo da autora). Para a autora, as formas em que não há a presença de um sujeito lexical,  $\emptyset$ +V3PS,  $\emptyset$ +V+SE e  $\emptyset$ +VINFIN são aquelas que expressam maiores graus de indeterminação, pois não são marcadas em relação à pessoa e, assim, apontam um distanciamento em relação às pessoas do discurso. Por outro lado, ainda segundo a autora, as formas *eu*, *você*, *a gente* e *nós* já não são tão genéricas, pois, mesmo não sendo possível identificar as pessoas, de algum modo, apresentam referências que são específicas e estão relacionadas aos agentes envolvidas no diálogo, assim, ela descreve que “o uso de *a gente* e *eu* implica obrigatoriamente na inclusão da 1ª pessoa, e *você* sempre envolve a 2ª (Milanez, 1982, p. 80, grifos da autora)

Embora o estudo de Milanez (1982) não seja um estudo sociolinguístico, ele é referenciado largamente por vários outros pesquisadores como Cunha (1993), Santana (2006), Dias (2017), entre outros, ao analisar o *grau* de indeterminação. Três situações para a indeterminação são previstas pela autora:

1. *Indeterminação completa* – quando não existe a possibilidade de recuperar o referente dentro de um contexto.

Doc.: E o senhor já foi ofendido?

Inf.: Não senhora, nunca não. Só de... desse bichinho que *Ø chama* la... lacraia.

Doc.: Ham.

Inf.: Que *Ø chama* candradão, isso aí já me ofendeu já. (Santana, 2006, p.89)

2. *Indeterminação parcial com referência implícita no contexto* – quando o referente não está presente no texto, nem se apresenta a partir de itens lexicais, mas, é possível depreendê-lo por meio de inferência.

Doc.: E vem cá, aqui tem muito morro?

Inf.: Possui bastante.

Doc.: Você já subiu em algum?

Inf.: Já. Eu já e muito. Quano **a gente** vai assim tirar casca, a gente sobe num morro arto. (Santana, 2006, p.89)

3. *Indeterminação parcial com referência explícita no contexto* – quando o referente não está presente no texto, mas, de alguma forma se apresenta a partir de itens lexicais e pode ser recuperado do contexto por meio de inferência.

Inf.: Mai no tempo que **a gente** era rapaz, que **a gente** juntava o grupo da gente todim... ah! O dia de hoje, nós tava hoje aqui? O forró tava aí por conta! (Santana, 2006, p.89)

Conforme Milanez (1982), a terceira situação proposta, a que se refere a uma indeterminação pronominal, aponta uma relação de dependência com o seu contexto, sendo assim “quanto mais a interpretação do pronome depende de elementos contextuais, mais próximo ele está de ser indeterminado” (Milanez, 1982, p. 43-44)

Desse modo, há um *continuum* de indeterminação, que vai do mais determinado (contexto em que o informante/emissor é incluído) para o menos indeterminado (contexto em que o sujeito discursivo não se encontra na composição do sujeito). Seguem alguns exemplos de Milanez (1982):

1. “*Falam que é difícil*”. (Mais indeterminado)
2. “Com o extrato que o banco manda, volta e meia **a gente** é surpreendido por algum engano. O banco dificilmente se enganará; pode ocorrer d`**a gente** de vez em quando errar na soma, na subtração” (Menos indeterminado) - (Milanez, 1982, p. 53).

Como o processo é complexo para identificar o grau de indeterminação do sujeito, é necessário observar com atenção os referentes antepostos e pospostos, ou seja, todas as pistas à disposição.

#### 5.4.2.1.2 Função da Indeterminação

Com referências em Milanez (1982), Pereira (2013) e Dias (2017), a variável função da indeterminação foi utilizada com o intuito de estabelecer as formas utilizadas a algumas das funções exercidas pela indeterminação e servem como hipóteses explicativas dos usos das formas de indeterminação. Das quatro funções categorizadas<sup>57</sup> (*desfocalização do sujeito*, *exemplificação*, *descomprometimento* e *ocultação do sujeito*), anteriormente por Milanez (1982), Pereira (2013) acrescentou mais uma função, a *economia linguística*, o que também foi utilizado por Dias (2017) em sua tese.

Sobre a primeira função, a *desfocalização do sujeito* em Dias (2017), ou *focalizar a ação verbal*, conforme Pereira (2013) – a estratégia de indeterminação é usada para tirar o foco do sujeito ou do referente. A ação verbal é focalizada e não o seu referente, como em: “mais deus mi livri di í numa festa aqui. . . eu tenhu até medu nem aniversáriu (a) **a genti** podi fazê aqui nessi lugá minina. (Pereira, 2013)”. A segunda função, a *Exemplificação* – é quando o falante faz referência a ele mesmo com o intuito de exemplificar uma situação bem genérica, comum a qualquer pessoa, a exemplo de: “si tem um rádiu mais altu (d) cê num consegui durmí (Pereira, 2013). Já a terceira função, conhecida como *descomprometimento* ou - “amenizar o efeito do uso da 1ª pessoa do singular” em Pereira (2013) corresponde ao fato do falante utilizar uma forma de indeterminação para não se comprometer na ação verbal mesmo que ele a tenha realizado, minimizando a sua responsabilidade. É exemplo dessa função em Pereira (2013) “ah (a) a genti vê qui EU vi não mais (b) a genti vê contá né? Assim...”

A quarta função estudada foi a *Ocultação do sujeito* ou – “esconder a identidade do referente do sujeito, conforme Pereira (2013) - é quando o falante utiliza de propósito a forma indeterminada para esconder a identidade do referente, como por exemplo: “a genti vê contá né? Assim... elis acharu uma muié aí né?”. Sobre essa função, Dias (2017, p. 128) ainda afirma que “trata-se de escamotear uma terceira pessoa, o que diferencia da função “descomprometimento””. Como essa função *ocultação do sujeito* pode ser confundida com o *descomprometimento*, Milanez (1982) reforça que:

---

<sup>57</sup> A nomenclatura utilizada será a mesma utilizada por Dias (2017).

[...] na última [descomprometimento] estariam em jogo os interesses do locutor em subtrair a sua própria pessoa do fato relatado, para não ter que assumir a responsabilidade de uma participação intencional no mesmo, enquanto que a ocultação ocorreria motivada pelo interesse do locutor em esconder, por alguma razão, a identidade de uma terceira pessoa (Milanez, 1982, p. 88, grifo da autora)

A última função, a *economia linguística*<sup>58</sup>, o falante não indica a identidade do referente mas ele dá algumas pistas de quem poderia ser ele, deixando para que o ouvinte deduza a identidade de referente, mesmo não tendo certeza de quem possa ser. Nesse contexto, têm-se uma indeterminação parcial que pode ser implícita ou explícita no contexto. Pereira (2013) exemplifica: “i foi na na sabesp. . . chegô lá pidiu a ligação (a) **elis** falô qui (b) **Ø** vinham vê si da:va pa fazê a ligação né?”

Conforme os autores, não é uma tarefa fácil identificar as funções de indeterminação, pois, é necessário analisar o contexto discursivo em que elas se encontram. São exemplos encontrados nessa amostra:

- (29) “Lá no município, **nós** tirávamos algumas arvores lá no município que tem, como é que se diz” (P17.F2M) - Desfocalização do sujeito
- (30) “[...] **nós** na caixa da cozinha sem saber o que fazer[...]” (P17.F2M) – Exemplificação
- (31) “[...]. Então cada dia que passa **nós** temos mais alimentos saudáveis [ ...]”. (P25.F1S)  
– Descomprometimento do sujeito
- (32) “Foi muito difícil. **Nós** vivemos quase um ano no *Cuito* e era insuportável”. (P29.F2S)  
- Ocultação do sujeito

#### 5.4.2.1.3 Tipo de frase

Ponte (2008), por exemplo, investigou essa variável, no entanto, ela não se mostrou produtiva pois as entrevistas realizadas objetivavam favorecer e incentivar somente o discurso do falante, diferentemente, de um diálogo. Como na entrevista há um incentivo pelo entrevistador para que o falante “discorra” sobre as situações apresentadas, os entrevistados produzirão mais textos declarativos (afirmativos e negativos, em especial, os primeiros) e, em menor grau, as interrogativas<sup>59</sup>. Neste estudo, partiu-se da hipótese de que os resultados seriam

<sup>58</sup> Não foram encontradas exemplos dessa variante na amostra.

<sup>59</sup> Não houve dados de frases interrogativas na amostra analisada.

similares, já que a natureza da entrevista foi parecida e teve o mesmo objetivo. São exemplos de frases encontradas no *corpus* de Angola:

(33) “**A gente** pensa numa maneira depois sai outra” (P15.M2M) - Frase afirmativa

(34) “[...] **a pessoa** não senta à vontade [...]” (P15.M2M) - Frase negativa

#### 5.4.2.1.4 Tipo de oração

Neste grupo de fatores propõe-se investigar se os tipos de oração exercem, de alguma maneira, influência nas escolhas pelas estratégias de indeterminação do sujeito pelos falantes luandenses. As orações a serem consideradas são quatro: absoluta, coordenada, principal e subordinada. A oração absoluta é aquela em que sua estrutura tem um elemento verbal ou locução verbal, constituída de períodos simples. A oração principal necessita de uma outra, neste caso, a oração subordinada, para atingir sua completude. Já, a oração coordenada, caracterizada por orações sintaticamente e semanticamente independentes.

Cunha (1993), Lopes (1993), Machado (1995), Carvalho (2010), Santana (2006), Assunção (2012), entre outros, investigaram esse grupo de fator e perceberam que tem uma certa relevância. Sobre as variáveis inovadoras, *a gente*, *nós* e *você*, pesquisas têm mostrado a exemplo de Cunha (1993) e Lopes (1993), em relação ao *nós* e *a gente*, que o pronome *nós* apresentou uma probabilidade em orações independentes enquanto o *a gente*, em orações dependentes. No que concerne a essa variável, pretende-se observar como ela se comporta no *corpus* investigado em Luanda:

(35) “[...] minha filha, **as pessoas** já não tem mesmo meios. [...]” (P09.F3B) - Oração absoluta

(36) “[...] **nós** não aprendíamos muito, mas aprendemos mais muito cantar.[...]” P09.F3B) - Oração coordenada

(37) “[...] **Você tem** que suportar porque tá a fazer o barulho[...]

(P29.F2S) - Oração principal

(38) [...], as coisas nem sempre são **conforme a gente** pensa. (P15.M2M) - Oração subordinada)

#### 5.4.2.1.5 Flexão verbal (SE)

Conforme Dias (2017), Nunes (1990) já tinha analisado um *corpus* do século XX com essa variável no falar paulista. O mesmo foi considerado por Dias (2017) para sua análise com textos baianos de períodos equivalentes (século XIX e XX). Para este grupo, as análises levadas em consideração se pautam na forma verbal com a presença da partícula *SE*, pois, muitas vezes, ela pode estar ligada com verbos flexionados como as formas do tempo finito. Portanto, há de se investigar também se as outras formas do modo infinitivo e gerúndio são capazes de influenciar no fenômeno nas diversas formas de indeterminar o sujeito pelos falantes em Luanda. Alguns exemplos são:

- (39) “Ø **nota-se** um crescimento muito grande da mulher a nível educacional, nas universidades” (P31.F3S) - Tempo finito
- (40) “Medo de morrer acho que todo mundo tem, né? Eu tenho, eu tenho medo de morrer porque quando **se falar** da morte... (P17.F2M) Tempo finito
- (41) “[...] “É uma evolução, não **se pode** permitir violência contra menores”. (P12.M1M) - Tempo infinitivo

#### 5.4.2.1.6 Tempo e modo verbal

Bechara (2019, p. 234) afirma que o nível temporal, o tempo verbal é aquele que “assinala a relação temporal do acontecimento comunicado como o momento do ato da fala; o presente encerra este momento, o passado é anterior e o futuro ocorrerá depois deste momento”. Sobre o modo, ele declara que é aquele que

[...] assinala a posição do falante com respeito à relação entre a ação verbal e seu agente ou fim, isto é, o que o falante pensa dessa relação. O falante pode considerar como algo feito, como verossímil - com um fato incerto -, como condicionada, como desejada pelo agente, como um ato que se exige do agente, etc. (Bechara, 2019, p. 234)

O autor classifica os modos verbais em indicativo, subjuntivo, condicional, optativo e imperativo. No entanto, nesta pesquisa, os modos são subdivididos nos três mais largamente utilizados e conhecidos pelos falantes. Assim, são considerados os seguintes fatores com alguns exemplos da amostra:

*Modo subjuntivo*<sup>60</sup>: presente, pretérito imperfeito e futuro.

*Modo imperativo*<sup>61</sup>: afirmativo e negativo.

<sup>60</sup> Não houve dados na amostra.

<sup>61</sup> Não houve dados na amostra.

*Modo indicativo*; presente, pretérito perfeito, pretérito imperfeito, futuro do presente, futuro do pretérito.

(42) “Cada tribo tem a sua língua. **Entendemos** que sou do quimbundo”. (P21. F3M) – presente

(43) “nós **corremos** o risco de poder contrair estas mesmas doenças” (P11.M1M) - pretérito perfeito

(44) “à tardinha **Ø brincávamos** a macaca, ringue, essas brincadeiras [...] (P07.F2B)

Os trabalhos de Menon (1994), Setti (1997), Godoy (1999) Santana, (2006), Assunção (2012) apontaram para o presente do indicativo como tempo/modo mais produtivo para a realização das estratégias de indeterminação encontradas em suas análises. Destaca-se que, em casos de ocorrências com locução verbal, foi considerado o tempo verbal do seu auxiliar, como atestaram Santana (2006), Carvalho (2010) e Assunção (2012).

Esta variável está ligada aos eventos fatuais (*reais/realis*) de referência e os não fatuais (*irreais/irrealis*). Nesse contexto, os *realis* estariam ligados aos fatos possíveis e hipotéticos, enquanto os *irrealis*, aos eventos não realizados. Os tempos presente e passado, por apontarem um evento fixo cuja ocorrência se dá em um tempo determinado, encontra-se na modalidade real, a exemplo do modo indicativo, enquanto o futuro, pelas ações não terem sido ainda realizadas, situações que ainda podem vir a ser, encontra-se na modalidade irreal, a exemplo do modo subjuntivo. Na pesquisa de Ponte (2008), as formas *nós* e  $\emptyset + V3PP$  foram favorecidas pelo modo *realis*, enquanto as formas *você* e *eles* pelo modo *irrealis*. Assim, nos dados da autora, embora a frequência maior tenha sido alcançada pelo modo *realis*, a forma pronominal *eles* favoreceu o modo *irrealis* com P.R. 0.6 contra P.R. 0.45 para os *realis*. Semelhantemente, essa situação pode ser, de alguma forma, visualizada nos dados de Assunção (2012). Embora ela não tenha controlado esse grupo de fatores (real/irreal), a maioria das ocorrências com a forma *você* aconteceu com o presente do indicativo, enquanto os resultados mostraram o futuro do subjuntivo como maiores favorecedores com P.R. 0.82 portanto, pode-se fazer uma relação, de alguma forma, com uma maior frequência para o modo *irrealis*. Esses resultados assemelham-se aos encontrados por Ponte (2008) quanto à variante *você*, por exemplo. Nesse aspecto, embora a análise aqui não tenha o foco de averiguar os modos *realis/irrealis*, os resultados para a variável tempo/modo verbal estão, de alguma forma, interligadas.

#### 5.4.2.1.7 Transitividade verbal

Sobre o grupo transitividade verbal, a análise é feita com base em três fatores. O primeiro, os verbos transitivos, são aqueles que necessitam de um complemento com ou sem intermédio de preposição. Em relação a esse fator, ressalta-se que não se pretendeu diferenciar na pesquisa se foram diretos ou indiretos. O segundo fator, os verbos intransitivos, que não necessitam de um argumento interno. O terceiro, os verbos copulativos (verbos de ligação). Seguem exemplos de cada um dos verbos, a seguir na amostra analisada:

(45) “[...] **nós** podemos **ser** assaltados, [...]”. (P15.M2M) - Verbo de ligação/copulativo

(46) “Muito. Muitos acidentes... quero dizer que **apanhou** o motorista distraidamente”  
(P21.F3M) - Verbo transitivo

(47) [...] a mulher **trabalhava** só para o homem. [...] (P30.M3S) - Verbo intransitivo

A hipótese a ser atestada nesse grupo de fatores é de que a variante verbo transitivo possa ser a mais favorecedora à regra de aplicação haja vista que, por apresentar um sentido de incompletude, é natural o uso de argumentos que preencham o sentido do verbo, e, conseqüentemente, da frase, sem falar que há ainda uma grande recorrência de usos de verbos transitivos na Língua Portuguesa pelos seus falantes.

#### 5.4.2.1.8 Tipo de verbo

Tamanine (2002) utilizou os dados do Projeto VARSUL/ Santa Catarina ao estudar o uso de *nós* e *a gente* nas cidades de Blumenau, Lages e Chapecó analisando o contexto tipo de verbo para observar se há influência desse fator na escolha por uma das variantes. A mesma autora, em 2010, também utilizou esse mesmo grupo de fatores para analisar os seus dados de Curitiba. Em suas pesquisas, Tamanine (2002, 2010) classificou os tipos de verbo em *dicendi*, de estado, de ação e epistêmicos. Os verbos *dicendi* são aqueles que estão relacionados com o dizer, que introduzem a fala e pensamentos, ou seja, o discurso. Alguns verbos que ilustram são: afirmar, dizer, perguntar, falar, contar, explicar, responder, conversar. Os verbos de estado são aqueles que descrevem um estado de ser, uma propriedade de uma situação ou condição do ser, e são exemplificados como ser, estar, tornar-se, virar, ficar, continuar, permanecer, entre outros. Os de ação, ao contrário dos de estado, expressam dinamicidade e ações e podem ser representados pelos verbos correr, comer, trabalhar, andar, fazer, etc. Já os verbos epistêmicos são aqueles

que estão relacionados à atividade mental e são exemplos os verbos pensar, lembrar, imaginar, achar, etc.

O mesmo contexto de Tamanine (2002) foi utilizado por Franceschini (2011) para justificar o uso da tipologia verbal, já que tem uma estreita relação com a escolha por uma das variantes. Além disso, foi também utilizado esse mesmo grupo de fatores aos contextos de variação envolvendo o *tu/você*. Em Tamanine (2002), por exemplo, os verbos *dicendi* foram os contextos mais favorecedores para o uso do *a gente*, com P.R. 0.57, enquanto os verbos de estado favoreceram o uso de *nós*, P.R. 0.61. Dados similares foram encontrados com os dados de Curitiba, em 2010, quando o *a gente* teve uma menor probabilidade relacionado aos verbos de estado com P.R. 0.32, seguido dos verbos *dicendi*, P.R. 0.36 e dos epistêmicos, 0.43.

Franceschini (2011) ao analisar os usos do *nós/a gente* e *tu/você* em Concórdia – SC na rodada realizada com os pronomes com referências indeterminadas, alguns grupos de fatores *tipo* e *tempo verbal* apresentaram-se significativos. Nessa pesquisa o pronome *a gente* apresentou uma maior probabilidade de uso com os verbos *dicendi* e epistêmicos, com P.R. 0.75 para ambos, enquanto o *nós* foi favorecido pelos verbos de estado e de ação, com P.R. 0.73 e 0.71, respectivamente.

Diante do que foi exposto, é interessante analisar mais essa estreita relação entre os tipos verbais e as formas para indeterminação do sujeito, principalmente, quando se investiga o *você*, *tu*, *nós* e *a gente* como recursos indeterminadores. A seguir, encontram-se exemplos de tipologia verbal encontrados na amostra de Luanda:

(48) “A gente temos que apertar mais e conversar com eles” (P10.F3B) -Verbo *dicendi*

(49) as dificuldades que **nós** vivemos fazem com que a gente dificulta mesmo criar nossos filhos nas condições necessárias. (P20.M3M) - Verbo de estado

(50) “**o levaram** meu negócio, onde é que eu meti em casa” - Verbo de ação

(51) “**Ø conseguimos** entender o que o outro fala, o Português é complicado.” (P15.M2M)  
-Verbo epistêmico

#### 5.4.2.1.9 Estrutura do núcleo do predicado

Dias (2017), ao tomar como base o estudo de Cavalcante (1999), verificou como as estruturas do núcleo do predicado podem ter uma influência no favorecimento de recursos de indeterminação utilizadas pelas pessoas, classificando-as em estruturas simples e complexas, no intuito de controlar os verbos auxiliares (tempo, modo e aspecto). Para as estruturas simples,

foi considerado a presença de um único verbo, enquanto, nas estruturas complexas, mais de um verbo.

Esse grupo de fatores não foi tão relevante nos dados encontrados por Dias (2017) já que a variável estrutura do núcleo do predicado foi selecionada apenas nas rodadas que envolviam o  $\emptyset+V+SE$  e o *a gente*, com prevalência das estruturas complexas. O autor faz um adendo dos seus dados quando verificada a frequência geral. Segundo ele, 84,7% das ocorrências acabaram sendo de estruturas simples e, desse modo, seus resultados assemelhavam com os encontrados por Cavalcante (1999). São exemplos encontrados no *corpus*:

(52) “aí **nós** fugimos” (P17. F2M) – Estruturas simples

(53) “E então vi, assim, que **estava fazendo** uns trabalhozinhos nas casas das senhoras pra aguentar estes negros filhos, e todos tavam comigo e eu mãe.” (P21. F3M) – Estruturas complexas

#### 5.4.2.1.10 Ausência ou presença de preposição

Menon (1994, p. 180), afirma que "no que concerne à indeterminação, a presença de variantes após as preposições é manifestada muito mais freqüente que se podia presumir". Setti (1997), com base nos estudos de Menon (1994), ao analisar os dados do NURC/SP, afirma que é importante observar as preposições com o infinitivo pessoal, pois, em muitos contextos, o preenchimento do sujeito tem ocorrido após as preposições que precedem de verbos que estão no infinitivo. Em muitos dos casos observados, o sujeito após uma preposição tem grandes possibilidades de ser genérico pois não é possível determinar seu agente. Setti (1997) ilustra essa situação em “Daí você não dá **tempo de** você ir no parque”.

Dias (2017) também considerou importante analisar a presença de preposição ligada a um verbo. Esse contexto foi verificado em seu estudo, a exemplo de “Entonce o cumpade é da minha pinião; nem é preciso pau pra se inducá fias.”. O autor, como base no estudo prévio de Cavalcante (1999), afirmou que, nessas situações, a preposição se liga, principalmente, aos verbos que se encontram no infinitivo sendo assim, os autores acreditam no favorecimento do *se* para a utilização de uma das formas para indeterminar o sujeito. Desse modo, é analisada nesta pesquisa a ausência ou presença da preposição<sup>62</sup> em seu papel de contribuição para a indeterminação do sujeito, sejam elas com o favorecimento do *se* ou de outra forma indeterminadora. Exemplo encontrado na amostra em relação à presença de preposição:

---

<sup>62</sup> Não houve ocorrência na amostra.

(54) “é uma gotícula d’água no oceano, não se **aprende**” (P22.M1S) – sem preposição

#### 5.4.2.1.11 Concordância com o argumento interno (partícula *SE*)

Castilho (2019, p. 480-481) afirma que os reflexivos estão perdendo o traço de pessoa, o traço de reflexividade, como também sofrendo com a perda no traço de apassivação. Sobre esse aspecto, ele cita a problemática do pronome apassivador e da voz passiva pronominal nas gramáticas e faz alguns questionamentos: “o verbo concorda ou não concorda com o substantivo plural? [...] Por que a concordância do verbo seguido de *se* está desaparecendo, juntamente com esse pronome?” Conforme o autor, baseado em Maurer Jr. (1951b), a explicação se encontraria na própria diacronia da língua, no latim vulgar, que agregou novas funções ao pronome *se* reflexivo, inclusive o de apassivador.

Castilho (2019) afirma que a perda de traços do pronome *se* provocou também algumas alterações na estrutura da passiva pronominal, como, por exemplo, o sentido passivo ficou comprometido e foi substituído pelo sentido de um sujeito indeterminado, como também desapareceu a concordância verbal com o sujeito na voz passiva, passando a ser reanalisado como objeto direto. De acordo com o autor, isso reflete a situação pela qual se encontra o PB – “a interpretação passiva (e, conseqüente concordância do verbo com o sintagma nominal no plural) se mantém apenas no estilo formal” (Castilho, 2019, p. 481). O autor ilustra essa situação com os seguintes exemplos:

- a) Vende-se flores. (alguém vende flores)
- b) Cortou-se os meninos. (= alguém cortou os meninos)

Nunes (1990) investigou o *se* por um viés sincrônico e diacrônico em construções com o *se* apassivador e o *se* indeterminador, em dois *corpora* orais e dois escritos e fez uma análise do grupo de fatores concordância com o argumento interno (partícula *SE*). A primeira variante se referia à concordância entre o verbo e o argumento interno, a exemplo de “Sendo medidas as ditas terras *se intregassem as quinhentas braças de terras* aos ditos Padres” e à discordância entre o verbo e o argumento interno “Também desenterrou-se *outras coisas semelhantes* em diversos lugares.”; além da neutralização da concordância, como em “As testemunhas depõem que se não deu cauza alguma aos índios Muras para fazerem os excessos referidos”. A segunda variável se referia à presença ou ausência do clítico “*se*”.

Em relação ao *se* indeterminador, a descrição de Nunes (1990), a seguir, converge com o que Castilho (2019) relata sobre a tendência cada vez mais da ausência de concordância ao pronome *se* no PB:

[...] o surgimento da passiva pronominal sem concordância [...] desencadeou um processo de mudança linguística no português brasileiro em que a forma conservadora (com concordância) foi paulatinamente perdendo terreno para sua concorrente, culminando no momento sincrônico, em que deve sua sobrevivência na modalidade escrita à renitência da gramática normativa. [...] no século XIX as passivas pronominais sem concordância superaram as passivas com concordância, tornando-se a forma canônica do século XX. (Nunes, 1990, p. 77)

Dias (2017) analisou a concordância com o argumento interno como um grupo de fatores em sua pesquisa. Nesse contexto, faz-se uma análise se o sujeito concorda com o verbo quando está relacionado com o pronome *se*. Embora as rodadas realizadas por Dias (2017) não apontaram relevância na seleção desse grupo de fator, seus dados apresentaram uma maior quantidade de concordância com o argumento interno do que a não concordância, independentemente da posição do argumento interno, diferindo dos dados de Cavalcanti (1999) que mostraram uma maior concordância do argumento interno na posição posposta, relativizado ou não realizado foneticamente.

Considerando que a tradição gramatical preza pela concordância do verbo com o sintagma nominal no plural, é interessante perceber como a comunidade de Luanda se comporta nesse contexto. Por isso, para esse grupo, serão consideradas as variantes de concordância com o argumento interno ou não-concordância quando o *se* está presente.

(55) “**Se chama** transporte em Angola, péssimo, por isso que há engarrafamentos”  
(P27.M2S) – Concordância com o argumento interno

(56) “Antes **considerava-se** as doenças sexualmente transmissíveis como um tabu”  
(P11.M1M) – Sem concordância com o argumento interno

#### 5.4.2.1.12 Posição do argumento interno do verbo

A posição do argumento interno do verbo também está relacionada ao *se*. O que se pretende observar aqui é se a posição anteposta ou posposta do SN pode interferir na concordância ou não do argumento interno, explicado no item anterior. Enquanto Berlinck (1988) e Nunes (1990) afirmam que existe uma maior probabilidade da não-concordância quando o sujeito é posposto. Tendo como base os autores citados acima, Dias (2017) esperava

encontrar uma maior quantidade de dados na posição posposta do SN (o que foi comprovado observando apenas os dados gerais, já que nenhum dos grupos de fatores ligados ao argumento interno do verbo foi selecionado estatisticamente), na posição relativizada ou zero foneticamente. É exemplo da comunidade de Luanda:

(57) “*Deveria se fazer* uma espécie dum inquérito”. (P22.M1S) - SN posposto

#### 5.4.2.1.13 Preenchimento do sujeito

Para este grupo de fatores, são analisadas duas situações: o preenchimento ou não do sujeito por um item lexical. Nesse contexto, para a variante do preenchimento, espera-se encontrar um núcleo não ocultado, ou seja, sujeito pronominal realizado, a exemplo da frase “*Eles* comeram muitas maçãs”, que seja aparente como previstas para alguns dos elementos marcados, a exemplo das variantes já destacadas em *nós*, *a gente*, *eles*, *você*, (cf. 4.4.1.2), ou outra estratégia que o falante do *corpus* possa ter utilizado. Já, as situações que existem a ausência de um núcleo do sujeito aparente, a não realização de um pronome, como em “ $\emptyset$  falaram bastante”, têm um sujeito não preenchido lexicalmente. Seguem exemplos para essas duas variantes:

(58) “agora **nós** temos é que tentar fazer os descontos” (P22.M1S) -Sujeito pronominal realizado)

(59) “ $\emptyset$  **tens** que ler bastante. Bastante mesmo” (P23.M1S) - Sujeito pronominal não realizado

#### 5.4.2.1.14 Mudança e Manutenção do referente (Paralelismo)

O paralelismo é considerado um fenômeno discursivo e marca a presença de repetições nos diversos níveis da língua como o fonológico, o morfológico, o lexical, o sintático e o semântico. Geralmente, o seu entendimento, no campo da Linguística, é tomado como uma tendência para a repetição de estruturas, de uma variante que já foi previamente usada.

O paralelismo é um elemento presente nas análises morfossintáticas e bastante relacionado com o uso de *nós* e *a gente*. Alguns estudos partem do pressuposto de que o paralelismo está associado a um discurso não planejado pelo falante ou considera que é uma

marca típica da oralidade. Sobre isso, Ochs (1979, p. 79) afirma que o fenômeno do paralelismo<sup>63</sup> que:

[...] it may be the case that, when speakers have not previously organized their discourse, they may retain the same morpho-syntactic format to express novel content. Hence, stream- of- consciousness writing and on-the-spot working out of a difficult concept may exhibit repeated use of a formal frame. (Ochs, 1979, p. 79)

pode ser o caso que, quando os falantes não têm o seu discurso organizado previamente, eles podem reter o mesmo padrão morfossintático para expressar o novo conteúdo. Assim, o registro do fluxo da consciência e a elaboração imediata de um conceito difícil podem exibir o uso repetido de uma estrutura formal. (Ochs, 1979, p. 79. Tradução nossa)

Já Tannen (1982) afirma que o fato de os falantes usarem as mesmas estruturas em seu discurso gera uma padronização rítmica fazendo com que os falantes não interrompam o seu turno de fala, sendo assim, essa estratégia pode ajudar os falantes a ganharem mais tempo enquanto pensam em uma informação nova.

Para Marcuschi (2006), o recurso da repetição é uma das estratégias mais utilizadas pelos falantes tem algumas finalidades. Dentre muitas das finalidades consideradas pelo autor, ele destaca que podem “contribuir para a organização discursiva e para a monitoração da coerência textual, favorecer a coesão e a geração de sequências mais compreensíveis, dar continuidade à organização tópica e auxiliar nas atividades interativas” (Marcuschi, 2006, p. 219). Para o autor, a repetição é uma estratégia central da oralidade, que está ligada a um novo conteúdo e que essas escolhas sintáticas estão intimamente ligadas com as situações de interação dentro do discurso. Sobre isso, Marcuschi (2006) levanta uma hipótese geral de que os padrões sintáticos na fala estariam em íntima relação com os padrões interacionais de tal forma que propriedades sintáticas de superfície são controladas nos níveis do discurso. A visão do autor corrobora os estudos funcionalistas de Givón (1979, p. 109), ao afirmar que “os princípios e parâmetros comunicativos que governam tanto o discurso como a sintaxe crescem do modus pragmático e de sua seleção ao longo da escala de registros da comunicação humana”.

Considerando o que foi discutido sobre o fenômeno do paralelismo e sobre estudos já previamente realizados sobre a indeterminação do sujeito como Santana (2006), Ponte (2008), Carvalho (2010) e Souza (2014), esta pesquisa também se propõe a analisar qual foi a forma

---

<sup>63</sup> Sobre o paralelismo, Ochs (1979, p. 79) afirma: “[...]pode ser o caso que, quando os falantes não têm o seu discurso organizado previamente, eles podem reter o mesmo padrão morfossintático para expressar o novo conteúdo. Assim, o registro do fluxo da consciência e a elaboração imediata de um conceito difícil podem exibir o uso repetido de uma estrutura formal. (Tradução nossa)

antecedente utilizada pelos falantes luandenses no discurso e qual variante foi escolhida por eles nas situações que se sucederam. O objetivo é observar se a forma utilizada anteriormente implicou na repetição de uma mesma forma ou se os participantes optaram por outra estratégia enquanto falavam com o/a entrevistador/a.

As variantes investigadas para mudança/manutenção do referente são: referente igual à estrutura anterior na série discursiva e o referente diferente da estrutura imediatamente anterior na série discursiva.

A hipótese que se busca investigar é que, numa cadeia discursiva, os falantes tenderiam a manter as marcas, de modo que as mesmas variantes seriam preservadas daquela verbalizada anteriormente. Acredita-se também que há mais probabilidade de se ter uma cadeia maior de repetições do que uma apenas.

(60) “**Ø** amarrávamos, **Ø** fazíamos nó e **Ø** secávamos a corda” (P17.F2M) – Manutenção do referente

(61) **Você** pode falar o Português, mas nunca **Ø** **vais** considerar que fala bem o Português, [...]”. (P15.M2M) - Mudança do referente

#### 5.4.2.1.15 Forma Antecedente (Paralelismo)

A variável forma antecedente tem sido bastante frequente nos estudos que tratam das estratégias que indeterminam o sujeito, não só pelo uso, mas também por apresentar grandes frequências no favorecimento dessas estratégias pois se liga a variável do Paralelismo (tendência a repetir as estruturas numa cadeia discursiva). Por isso, com o objetivo de identificar melhor os seus antecedentes, é muito importante investigar se a primeira forma apresentada pelo falante, fará com que ele utilize o mesmo recurso. Assim, é possível perceber se os falantes manterão ou não, no decorrer da sentença, a estratégia que eles usaram inicialmente quando o seu referente é mantido. Utilizaram-se as formas<sup>64</sup> usadas por Guimarães (2018), a saber: forma precedida de a gente explícito, forma precedida de a gente implícito, forma precedida de eles explícito, Forma precedida de eles implícito, forma precedida de eu explícito, forma precedida de eu implícito, forma precedida de FN explícito, forma precedida de FN implícito, forma precedida de nós explícito, forma precedida de nós implícito, forma precedida de você explícito, forma precedida de você implícito, forma precedida de Ø+V+SE, forma precedida de Ø+V3PP,

<sup>64</sup> Não houve casos de formas precedidas para as seguintes variantes: de “eu explícito” e “eu implícito”, de “forma nominal implícita”, “voz passiva analítica”, “voz passiva sintética” e “Ø+V2PS”;

forma precedida de Ø+V3PS, forma precedida de Ø+V1PP, forma precedida de Ø+VINFIN). Seguem alguns exemplos encontrados para alguns dos fatores:

- (62) “quando você gostar duma miúda você primeiro tem que falar ainda - forma precedida de você explícito (P01.M1B)
- (63) “[...] os clientes não aparece. Assim conforme que se vê aqui, conforme Ø tamos no armazém [...]” (P01.M1B) - forma precedida de Ø+V+SE
- (64) “Logo ø viram que as coisas foi muito assim grave no dava pra eles socorrer. (P01.M1B) - forma precedida de ØV3PP
- (65) “[...] então **eles** vão domingo e só voltam na sexta-feira. [...]” (P15.M2M) – forma precedida por ele explícito

#### 5.4.2.1.16 Inclusão do Falante

Esse grupo de fator foi investigado por Ponte (2008) e a inclusão do falante se dá porque, muitas vezes, o falante faz uma referência a um sujeito indeterminado com a possibilidade dele mesmo ser incluído na ação. Ponte (2008) obteve resultados significativos para esse grupo quando a indeterminação do sujeito estava ligada às estratégias de *você* e *ele* e apontou, por um lado, o fato de o falante raramente se excluir quando utilizava o pronome *você* e, por outro lado, ao usar o pronome *eles*, o fato da não-inclusão do falante foi uma marca categórica. Desse modo, convém observar, dentre os recursos utilizados pelos entrevistados de Luanda para indeterminar o sujeito, se houve a inclusão ou não do falante no universo da referência realizada por ele. Assim as variáveis utilizadas serão a inclusão e a não inclusão do falante. Exemplos encontrados na amostra pesquisada são:

- (66) “E se **repararmos** o que acontece na...no pessoal que vive nas periferias – (P.22.M1S) - Inclusão do falante
- (67) “Mas só que não **ø encontrou** nada. Não encontrou nada em casa”. (P.21.F1B) - Não inclusão do falante

#### 5.4.2.2 Variáveis Extralinguísticas

Na subseção 5.4.2.1 foram apresentadas as variáveis independentes, as linguísticas, e nesta subseção, são apresentadas as variáveis extralinguísticas que estão ligadas diretamente ao

*corpus* investigado e que contribuem para o entendimento do comportamento linguístico pelos falantes em relação ao fenômeno da indeterminação do sujeito na comunidade de Luanda. Portanto, as variáveis selecionadas são: sexo, faixa etária e escolaridade.

#### 5.4.2.2.1 Sexo

Nos estudos sociolinguísticos, é apontado que as mulheres tendem a liderar os processos de mudança linguística. Desse modo, costumam empregar as formas inovadoras quando não há estigma social.

Trudgill (2000) afirma que a diferenciação de sexo surge porque a língua, como um fenômeno social, está relacionada às atitudes sociais dos falantes. Assim,

Men and women are socially different in that society lays down different social roles for them and expects different behaviour patterns from them. Language simply reflects this social fact. If the social roles of man and women change, moreover, as they seem to be doing currently in many societies, then it is likely that gender differences in language will change or diminish also, and we have seen above some evidence that this may be so. (Trudgill, 2000, p. 79)<sup>65</sup>

Conforme Trudgill (1979, *apud* Monteiro, 2000, p. 73-74), em sua pesquisa sobre o inglês falado na Inglaterra, o fato de as mulheres usarem mais as formas de prestígio com mais frequência que os homens se explicaria, basicamente, por alguns motivos. O primeiro deles, a “posição subordinada das mulheres na sociedade exige delas tal comportamento, para que assegurem, pelo menos o *status* linguístico” e, o segundo, se refere ao fato de que os homens “podem ser avaliados socialmente pelo que fazem, as mulheres são avaliadas principalmente pelo que aparentam”. As mulheres, ao longo dos anos, têm lutado, não só no Brasil, mas também em Angola para que não continuem sendo excluídas, marginalizadas e vítimas de uma sociedade patriarcal e extremamente machista. As consequências vão além da busca pelo *status* linguístico, pela cidadania, pelos empregos e salários justos, pelos sucessos nos estudos e cargos de liderança, mas, principalmente, reparação social e empoderamento.

Há também de se considerar, portanto, que as diferenças no uso da língua entre homens e mulheres vão além de questões biológicas como o timbre, altura da voz, por exemplo, como

---

<sup>65</sup> Homens e mulheres são socialmente diferentes no sentido de que a sociedade lhes confere papéis sociais distintos e espera padrões diferentes de comportamento. A língua simplesmente reflete o fato social. Os papéis sociais dos homens e mulheres mudam em muitas sociedades, então é provável que as diferenças de gênero na língua poderão mudar ou diminuir também, e já se tem visto evidências de que isso possa ser possível. (Tradução nossa)

refletem também diferenças nos papéis sociais que cada um desempenha na sociedade como na forma que elas se socializam dentro da comunidade. Labov (2008 [1972], p. 348) afirma sobre o sexo que “A diferenciação sexual dos falantes não é, portanto, somente um produto de fatores físicos, ou de diferentes quantidades de informação referencial fornecida por eles, mas, sim, uma postura expressiva que é socialmente mais apropriada para um sexo do que para outro.”

Paiva (2010, p. 35) reitera que a “análise da correlação entre gênero/sexo e a variação linguística tem de, necessariamente, fazer referência não só ao prestígio atribuído pela comunidade às variantes linguísticas como também a forma de organização social de uma dada comunidade de fala”. Inevitavelmente, as pressões sociais existem sobre os falantes para que usem as formas de prestígio, que são as da classe dominante. Se, por um lado, as mulheres tendem a implementarem na língua uma forma que é mais socialmente prestigiada, por outro lado, os homens são aqueles que encabeçam a implementação de usos socialmente desprestigiados. Desse modo, o emprego das formas inovadoras tem aparecido com mais frequência entre as mulheres do que os homens quando desprovidas de estigma social, como também, elas costumam empregar mais que os homens as formas que as gramáticas normativas prescrevem como próprias da língua-padrão.

Conforme Labov 2008 [1972], p. 346),

Na medida em que os pais influenciam a língua inicial das crianças, as mulheres influenciam mais ainda; as mulheres certamente conversam mais do que os homens com as criancinhas e têm uma influência mais direta durante os anos em que as crianças estão formando regras linguísticas com maior rapidez e eficiência. Parece provável que o ritmo do progresso e a direção da mudança linguística devem muito à especial sensibilidade das mulheres a todo o processo.

Ainda, conforme o autor supracitado, não se deve generalizar a ideia de que as mulheres são as que sempre encabeçam a mudança linguística, pois alguns estudos realizados por ele mesmo na Ilha de Martha’s Vineyard, mostram indícios de que o inverso também pode acontecer, já que os homens falam mais “com a boca fechada”, além de utilizarem áreas mais contraídas da fonologia do que as mulheres. Um outro exemplo que pode ser tomado aqui para essa generalização equívoca é o estudo de Omena (1996, p. 14 *apud* Paiva, 2010, p. 36) sobre o uso do *nós* (variante mais conservadora) e *a gente* (variante mais inovadora), visto que a variante inovadora foi mais usada entre as mulheres, entre os mais velhos, visto que “as duas variantes não se sujeitam a uma avaliação social explícita ou à exclusão normativa”.

Quanto à variável sexo, nesse estudo, a hipótese é que tanto homens e mulheres utilizem as distintas formas de indeterminar o sujeito que não só aquelas previstas pela norma padrão,

no entanto, com valores maiores para as mulheres, acreditando que esse fenômeno não é estigmatizado pela comunidade pesquisada.

#### 5.4.2.2.2 Faixa Etária

A faixa etária é uma variável interessante para se analisar em pesquisas linguísticas. Ao tentar avaliar o comportamento que os falantes têm, não só analisando a sua idade real como também pelos intervalos de idades em que se encontram, é possível prever se determinado fenômeno em questão que se propõe investigar pode ser um indicativo de uma variação, onde as variantes, simplesmente, coexistem, se é uma mudança em curso ou até mesmo uma mudança implementada na comunidade.

Tanto os mais jovens quanto os mais velhos não falam do mesmo jeito. Alguns falantes até podem mudar algum hábito linguístico em algum momento da vida, mas a comunidade como um todo pode não modificar, não acarretando assim, mudança na língua. A idade do falante, uma variável contínua, atrelada as outras variáveis ajudam na compreensão do comportamento linguístico na comunidade, por isso, nesta pesquisa levou-se em conta a faixa etária a que o indivíduo pertence.

O estudo é de tempo aparente e com uma análise feita em uma perspectiva sincrônica. Nesse tipo de estudo é possível observar se há indícios de uma mudança em curso ou não. Desse modo, a análise dos dados pode apontar como o uso das formas mais inovadoras estão sendo ou não utilizadas pelos falantes luandenses e com quais frequências. Vale ressaltar aqui que, se os dados observados indicarem mudança em progresso, só um outro estudo em tempo real, perspectiva diacrônica, de longa ou curta duração, poderá dizer com maior precisão a real situação linguística da comunidade investigada.

As faixas etárias foram esquematizadas da seguinte maneira, conforme previamente apresentado: Faixa I (21 a 35), Faixa II (36 a 51) e Faixa III (acima de 52 anos). Normalmente, os falantes mais velhos preferem usar as formas mais conservadoras, e os mais jovens, as mais inovadoras favorecendo a aplicação da regra das formas alternativas de indeterminação do sujeito. A escolaridade, de alguma forma, está entrelaçada com a faixa etária. Nas pesquisas sociolinguísticas, muitos resultados apontam que, quanto mais tempo as pessoas passam na escola, a tendência é utilizar as formas padrão enquanto os menos escolarizados, as formas mais inovadoras. Espera-se encontrar uma curva ascendente quanto ao uso das formas aceitas pelas gramáticas à medida que os indivíduos são mais velhos e escolarizados.

Considerando que nem todos os falantes tiveram acesso à escolarização, havendo inclusive, participantes que não frequentaram a escola ou que tiveram uma instrução muito insipiente, acredita-se que, independentemente da idade, devido a sócio-história da comunidade, os falantes vão utilizar as formas mais inovadoras, pois não é possível precisar se todos os falantes entraram no período regular na escola. Alguns podem ter frequentado a escola mais tardiamente ou, até mesmo, nem frequentado. Por outro lado, acredita-se que os falantes que tiveram mais tempo no ambiente escolar podem utilizar mais a forma padrão devido à ação do processo de escolarização.

#### 5.4.2.2.3 Escolaridade

O espaço escolar deve ser o local de promoção e ampliação do repertório linguístico do falante, o espaço escolhido para que as pessoas possam adquirir uma competência comunicativa mais ampla e diversificada que vá além das formas já aprendidas em ambiente familiar e com os amigos. A distribuição dos bens culturais deve ser justa e a escola é um dos espaços em que esse direito democrático seja garantido. Sobre a função social que as instituições de ensino devem assegurar, Bortoni-Ricardo (2004, p. 75) afirma que:

*Ao chegar à escola, a criança, o jovem ou o adulto já são usuários competentes de sua língua materna, mas têm de ampliar a gama de seus recursos comunicativos para poder atender às convenções sociais, que definem o uso linguístico adequado a cada gênero textual, a cada tarefa comunicativa, a cada tipo de interação. Os usos da língua são práticas sociais e muitas delas são extremamente especializadas, isto é, exigem vocabulário específico e formações sintáticas que estão abonadas nas gramáticas normativas. (grifo da autora)*

No entanto, o ambiente escolar tem se tornado cada vez mais um local que está preso à ideia de “certo” e “errado”, tanto na modalidade oral quanto escrita. A escola pode até considerar formas distintas usadas pelos falantes como legítimas, mas não deixa de defender a forma de prestígio. Desse modo, a escola tem sido responsável pelo desrespeito às variedades linguísticas, sobretudo, aquelas mais relacionadas às formas sem prestígio e até das diversas línguas minorizadas como é o caso do ensino em Angola. No caso do país africano, a situação linguística é ainda mais complicada e delicada, visto a grande quantidade de outras línguas em contato com a Língua Portuguesa no território. Em sociedades que se encontram em um mosaico multilíngue como é o caso da sociedade angolana, a complexidade deveria ser vista de um ponto de vista positivo, no entanto, a questão crucial é saber como a escola lida com esse

complexo e instigante mosaico. Não se pode também perder de vista como a formação de professores para lidar com essas questões tão profundas estão de fato acontecendo.

A escola, diante de certas variantes dentro de uma comunidade, assume uma postura conservadora e sempre se justifica em um ideal de variedade padrão como um modelo ideal de língua a ser seguido pelos falantes. Desse modo, ela atua a favor de uma pureza da língua, e luta constantemente contra os “erros” ou “vícios de linguagem”. As formas mais estigmatizadas, desprestigiadas, moldam um falante que é ignorante e que “corrói” a língua.

Votre (2010, p. 51) afirma que a escola:

[...] gera mudanças na fala e na escrita das pessoas que as frequentam e das comunidades discursivas. Constata-se, por outro lado, que ela atua como preservadora de formas de prestígio, face a tendências de mudança em curso nessas comunidades [...] a escola incute gostos, normas, padrões estéticos e morais em face da conformidade de dizer e de escrever. (Votre, 2010, p. 51)

A influência da variável nível de escolaridade é muito importante para compreender como a escola influencia no repertório sociolinguístico do falante. Acredita-se que não só a quantidade de anos que um indivíduo passa no ambiente escolar, como também a qualidade das escolas que ele frequenta atua de modo a preservar as formas de prestígio face à tendência de mudança. Em Moçambique, por exemplo, os alunos têm muitas dificuldades na aprendizagem da língua portuguesa, não só por falantes de L2 e, esse fato é agravado pela fraca formação dos professores. Os alunos que não dominam a LP estão fadados ao insucesso em várias outras disciplinas. Como se não bastasse, a dificuldade dos professores, e conseqüentemente, dos alunos, os manuais escolares ainda reproduzem nos espaços escolares a Norma Padrão Europeia, enquanto, as línguas bantu não tem espaço nas escolas. O ensino ainda é bem tradicional que, ao buscar uma padronização, uma homogeneidade linguística na formação dos alunos, acaba elevando ainda mais os níveis de reprovação e ineficiência do aprendizado de LP e dos demais componentes curriculares.

Deve-se analisar que o grau de escolarização de um indivíduo está intimamente ligado ao *status* socioeconômico, pois o preconceito linguístico é ainda mais forte se a forma utilizada da língua acontece com uma classe que é socialmente discriminada. Considerando que, nem no Brasil e nem em Angola, há uma distribuição de renda e de bens materiais equitativos, é muito provável que essas diferenças linguísticas sejam também mais evidenciadas. Outro ponto em comum entre os dois países citados é que, muitas vezes, as classes menos desfavorecidas são aquelas que tem índices de analfabetismo maiores. Vale ressaltar que o inverso pode não ser

verdadeiro, já que nem sempre, nas comunidades que possuem falantes com nível de escolaridade mais avançado, como o ensino superior, não, necessariamente, fazem parte de estratos mais elevados da sociedade.

Espera-se que a variação linguística implique em uma variação, sobretudo, social, cujas variantes empregadas por aqueles considerados de classe socioeconômica mais baixa da população, são, em sua maioria, desprestigiadas e estigmatizadas. Muitas vezes, na tentativa de frear as variantes inovadoras que naturalmente vão surgindo, os falantes, ainda, podem acabar utilizando outras formas alternativas, estratégias de esquivas, menos artificiais.

Labov (2008[1972]) reforça que as estruturas linguísticas refletem os processos sociais, pois existe um espectro de benefícios nas pesquisas que envolvem sociedade e a língua, sendo um deles:

Uma vez determinada a significação social de dada variante linguística [...] esta variável pode então servir de indicador para medir outras formas de comportamento social: aspirações à ascensão social, mobilidade e insegurança social, mudanças na estratificação e segregação social. [...] A comunidade de fala não é definida por nenhuma concordância marcada pelo uso dos elementos linguísticos, mas sim pela participação num conjunto de normas compartilhadas. (Labov, 2008[1972], p. 150)

O grau de instrução do falante pode condicioná-lo ao uso de determinada(s) forma(s) linguística(s). Acredita-se que, os falantes que ficaram mais tempo no ambiente escolar tiveram mais acessos aos bens culturais, e portanto, devem favorecer as formas de prestígio.

Desse modo, espera-se que o pronome *se* como indeterminador do sujeito, possivelmente, não esteja tão presente na fala daqueles com graus de escolaridade mais baixos, visto que essa é uma estratégia aprendida na escola, assim como as variantes mais inovadoras sejam mais utilizadas por falantes menos escolarizados. Assim, à medida que os anos de escolarização vão aumentando, espera-se daqueles que tenham frequentado mais tempo os espaços escolares, produzirão um maior número de dados da variante considerada padrão.

Na seção seguinte, serão feitas as análises após os dados codificados e quantificados em ferramenta computacional com base no *corpus* investigado, tecendo comparações com os estudos realizados sobre as estratégias de indeterminação que os falantes no Português de Luanda, capital de Angola.

#### 5.4.2.2.4 Língua de Intercomunicação

Considerando que muitos angolanos estão inseridos em meio ao bilinguismo por razões sócio-históricas, que já foram apresentadas na seção 2, é interessante analisar aqui, se os Luandenses que foram entrevistados neste *corpus* pudessem oferecer algum outro resultado que esteva ligado a situações de bilinguismo que é tão comum e tão presente nas vidas deles, mesmo sendo falantes do português como L1. O objetivo foi compreender qual grupo de falantes tendem a utilizar formas mais genéricas para o sujeito. As variantes analisadas para este grupo foram: “fala apenas português”, falante monolíngue e “fala português e compreende/fala língua bantu”, bilíngues.

Na seção seguinte, são apresentados e discutidos os resultados dos dados gerados pelo programa estatístico, o *Goldvarb X*, a fim de traçar respostas sobre o fenômeno variável em questão por meio de fatores linguísticos e extralinguísticos.

## 6 AS ANÁLISES E INTERPRETAÇÕES SOCIAIS E LINGUÍSTICAS DOS DADOS DE FALA DO POVO DE LUANDA

Nesta seção, são apresentados os dados com base nas variáveis descritas para a investigação na seção 5, os resultados das análises e as suas interpretações, tendo como base as variáveis selecionadas e apontadas pelo programa como mais favorecedoras. São feitas análises com vistas a conhecer a real situação linguística da cidade de Luanda em relação ao tema desse estudo. Confrontam-se as variáveis dependentes mais favorecedoras com outras que também foram consideradas importantes, além de possíveis realizações de análises binárias. São mostradas as variáveis que também não se apresentaram favorecedoras neste estudo com o objetivo de confrontar/comparar resultados de outras pesquisas. É nesta seção que são descritas as estratégias que mais foram utilizadas e as que menos contribuíram para a análise na comunidade investigada. Valores percentuais e de pesos relativos serão tomados como parâmetros para tal análise.

Antes dos resultados estatísticos obtidos por meio do programa *Goldvarb X* serem apresentados, faz-se um breve panorama das principais questões sociais observadas nas falas do povo luandense, as vozes de Luanda. As ricas e sofridas histórias de vida contadas por cada um dos participantes nas entrevistas analisadas exemplificam muitas das questões sócio-históricas enfrentadas pelo seu povo, suas tradições, suas percepções em relação à sua língua autóctone e à língua portuguesa, os preconceitos atrelados aos aspectos socioeconômicos e linguísticos, já que também há escolhas por determinados termos, estratégias linguísticas utilizadas pelos luandenses.

### 6.1 VOZES DE LUANDA EM ANÁLISE: O SOCIAL E LINGUÍSTICO SE ENCONTRAM

Nesta seção, são apresentadas experiências vividas pelos participantes em Angola. Muitos aspectos sociais são abordados, explorando suas perspectivas sobre uma variedade de tópicos. Como já mencionado em seção anterior, as entrevistas abordam temas como infância, educação, trabalho, transporte, vida familiar, relações interpessoais e costumes. A maioria dos informantes discute como a vida era diferente no passado em relação à atualidade em Angola, com foco particular nas mudanças sociais e culturais.

Há também uma tristeza silenciosa pelo que se perde da própria identidade, misturada às influências que vêm de fora, vêm do Brasil e de tantos outros lugares, e acabam “apagando”

algumas marcas culturais. O tom das entrevistas é informal e reflete as experiências pessoais e alguns dos aspectos linguísticos dos indivíduos que ora os aproximam, ora os distanciam.

As brincadeiras infantis refletem a riqueza cultural e são essenciais para a formação social. Geralmente, elas são passadas de geração em geração, preservam e transmitem os valores, a tradição e os saberes populares. Elas também podem ser utilizadas como um modo de ensinar e integrar a cultura local, promovendo ao mesmo tempo a interação e o desenvolvimento de habilidades.

Os entrevistados revivem as brincadeiras que preenchiam seus dias, lembrando o “Agriã”, a “Barra de lenço”, “Macaca” e “Ringue”, com pedrinhas, bonecas. As brincadeiras tinham regras e disputas simples, nas quais até um pedaço de tecido se tornava um troféu, um motivo para competir e, ao mesmo tempo, se unir.

(68) “Eu não sei bem porque mas antigamente o único dia que nós brincávamos, -os pais dava aquelas leis - só aos domingos íamos à igreja fazíamos o serviço todo de casa, à tardinha brincávamos a macaca, ringue, essas brincadeiras escondidas e mais outras brincadeiras.(P07.F2B)

Havia também a pesca com anzol e o trabalho na lavoura, onde o esforço se misturava ao aprendizado, e cada gesto era uma lição de vida para os moradores de Luanda, como relata um motorista veterano que guia sua vida entre as ruas e os destinos dos outros. Quando a voz do rádio se cala, ele relembra um pouco das suas novelas favoritas e descreve as dificuldades de criar filhos no Huambo, onde o mundo parece duro e o passado distante. Entre uma corrida e outra, ele divaga sobre o amor e a morte, parceiros inseparáveis de suas reflexões diárias.

(69) “As nossas brincadeiras que nós fazíamos nós aprendíamos de quase muito.... A nossa brincadeira era mais muito se ensinar a pesca no rio. Então a gente vamos no rio, pescar, voltamos ao rio umas horas que e tal é prontos brincávamos mais muito com a bola, jogar a bola, Certo.” (P01.M1B)

O amor nascia tímido, guardado nas sombras da tradição e sob o olhar atento das famílias. O namoro era discreto, um compromisso silencioso, respeitoso, no qual o encontro era um ato reservado, e a noiva surgia ao público apenas no grande dia, envolta em mistério. Diversos entrevistados comparam o namoro no passado com o presente, notando diferenças significativas e descrevendo o namoro tradicional. Alguns destacam que uma das maiores diferenças é que o namoro atual é muitas vezes baseado em interesses financeiros e sexo. Outros contrastam o namoro em Luanda com o do Huambo, criticando a falta de envolvimento familiar, a

superficialidade e a precocidade dos relacionamentos atuais. Agora, o amor é bem diferente, ele se mostra rápido e frágil, perdendo-se na superfície, em conveniências, num ritmo que deixa de lado o respeito às raízes e o lugar das famílias. Há um lamento pelo que se foi, pelo que já não se guarda com o mesmo valor.

(70) “Quer dizer, o namoro daqui o namoro do Huambo é muito, muito diferente. Porque o namoro daqui pra pessoa ainda pode namorar, com miúda, enquanto que nem em casa, na casa da miúda nem sabes e nem sabes dizer que a nossa filha esta a andar com o fulano, mas enquanto que ela afinal já conhece que fulano que tá andar com ele.” (P01.M1B)

O transporte público em Luanda é um aspecto essencial do cotidiano. Nos inquéritos, percebe-se que a população enfrenta vários desafios, principalmente em relação aos táxis. Chamados de candongueiros, geralmente são vans ou carros maiores. São acessíveis, mas com sérios problemas relativos à segurança, como o excesso de velocidade, desrespeito às normas de trânsito e aos passageiros com o som nas alturas. Uma entrevistada, mulher de 54 anos, enfrentou a vida analfabeta agarrando-se a cada letra quando teve a chance, bem depois da infância, no esforço das aulas de alfabetização para adultos. Hoje, ela se mostra espantada com a falta de respeito:

(71) “os próprios taxistas não têm maneira de como andar na rua. Primeiro já que não respeitam os passageiros, levam os passageiros de uma forma imprópria, é... daí o pessoal manda parar o táxi, antes desta passageira ou passageiro subir e assim apanhar o lugar, ele arranca com uma velocidade e num diz nada” (P21.F2C)

Os candongueiros transportam as histórias da cidade, mas, para muitos, são símbolo de desconforto e superlotação. Os entrevistados falam da música alta, das corridas tumultuadas, e do comportamento dos motoristas, refletindo uma luta diária entre o necessário e o caótico.

(72) “o sistema de transporte cá é complicado! É muito complicado, é um Deus que nos acuda. Apertos de um lado, apertos de outro... a pessoa não senta à vontade, enfim (...) é muito [atabalhado] mesmo.” (P15.M2B)

Em Luanda, as marcas do presente se desenham também na escassez. A pobreza e as oportunidades limitadas são temas recorrentes, onde o alto custo de vida transforma cada conquista em um desafio, e o dia a dia exige esforço para se equilibrar entre o querer e o poder.

A guerra civil se manifesta em fragmentos de memórias, nas entrelinhas das histórias contadas. A experiência da guerra é individual, marcando cada entrevistado de maneira única.

A dor da perda e a incerteza do destino são algumas feridas abertas que moldam as visões de mundo. Um evento traumático que os obrigou a fugir de suas casas e que também se manifesta na consciência da dificuldade de criar filhos em um ambiente marcado pela violência e pela pobreza.

(73) “Ele foi trabalhar lá mesmo lá no Huambo, lá no contorno do Huambo e pronto lá já havia guerra e vieram dizer na minha mãe que meu pai não aparece lá passou uma guerra, vieram dizer minha mãe que meu pai não apareceu e onde, no se sabe, no se sabe explicar se morreu ou tá em vida, até agora. Mas em noventa e dois quando veio a paz cá, cá, cá em Angola muita gente que vieram mas o meu pai no, no ate agora no consi...nunca ouvimos nada dele. A minha mãe também morreu de guerra, assim que eles atacaram, quem sabe explicar só é o meu irmão que sabe. Já no lembro. O meu irmão diz que a minha mãe foram trocar sei lá o que, comida, tiraram as coisas vieram dizer minha mãe que meu pai não apareceu e onde, no se sabe, no se sabe explicar se morreu ou tá em vida, até agora”. (P03.F1B)

(74) “A guerra que a gente sofreu é a guerra desse Massacre do (...). Massacre do (...) eu vivia no Golfo. Então (...), eu vou visitar em fim de semana vou visitar uma comadre, era...nesse foi comadre do falecido meu tio. Vou visitar no Sambizangae este dia que eu vou no Sambizanga, que eu dormi, na mesma noite arrebentou a guerra. Pronto, a gente era, era... se a gente não morremo porque já enfim memo, a gente... Deus tava conosco. A gente corria naquela noite, naqueles lado ali da linha férrea embaixo (...) outro caiu à frente, outro caiu atrás, você pula só Jesus mesmo é que sabe. Dali pronto (...). Fomo para a Funda. Funda a gente ficamos nos aviários. Havia uns aviários aí aonde nós ficamos e o (...) quando ia à frente (...). Ficamo aí nos aviário era só comer galinha e galinha, e ultimamente as galinha eram os (...). Galinha (...) milho, mata bicho galinha, almoço é galinha, jantar é galinha, ficamos aí no Massacre do (...), mas aquilo era mortes de verdade. (...), até quando Deus nos ajudou, a guerra cessou, saímos daí dos aviários chegamos aqui em Luanda, as casas todo, encontramos as casas todo (...) sabotara, muitas mortes. Pronto. Aquilo foi um sofrimento memo de andar à pé daqui de Luanda até na Funda é grande sofrimento. Aquilo se a gente... os outros ficaram daí pronto, Deus nos ajudou, nos salvamos e os outros ficaram. Aquilo não era (...), aquilo era mosca. Massacre do (...), aquilo mesmo, se a gente não morremo porque já, enfim...” (P10.F3B)

Embora a guerra seja uma presença constante, os textos também revelam a resiliência do povo angolano. As histórias de vida demonstram a capacidade de superar as adversidades, de reconstruir a vida em meio ao caos, de encontrar alegria e esperança. Mesmo em tempos difíceis, são vozes que convidam a refletir sobre o impacto da guerra na vida das pessoas e a importância de construir um futuro de paz e prosperidade.

A vida em Luanda é, frequentemente, descrita como difícil. Os entrevistados relatam as dificuldades de criar seus filhos na capital, tendo que trabalhar em casas de outras pessoas para sobreviver. Luanda é também descrita como uma cidade agitada e estressante, com muita contrariedade, um lugar para "desenrascar" e juntar dinheiro para enviar à família no interior.

O saber não era para todos, e a escola era um privilégio raro. Muitos encararam a dificuldade do acesso à educação formal, enquanto outros mergulharam nas lições das missões

religiosas, onde o ensino vinha misturado com os preceitos da fé, como um norte espiritual. A pobreza muitas vezes limitava seus moradores às oportunidades educacionais e o ciclo da desigualdade social perpetuava.

(75) “tia, quero ir também ir para a escola”  
 “ir à escola fazer o que? Achas que eu capacidade de pagar a escola para ti?” (P04.F1B)

A preservação das línguas nacionais é defendida por diversos entrevistados, e eles destacam a necessidade de valorizar e ensinar as línguas nacionais às novas gerações que é de grande relevância para a identidade cultural. A dificuldade com a língua portuguesa também era um tema recorrente. Em Luanda, é destacada uma particularidade bem característica do povo, com uma identidade própria falada por seus moradores. Mesmo exaltando as línguas nacionais, os entrevistados destacam a importância do português como língua comum para a comunicação entre pessoas de diferentes províncias. Essa questão é muito bem discutida por Silva e Araujo (2020).

(76) “Praticamente eu digo assim já tamos acostumado, nós batalhámos memo pra educar os filhos, mas as vezes o Deus dá mesmo muita força, principalmente nós que no estudamos eu vejo que a educação dos nossos filhos são muito deferente e são filhos que não dão assim muito trabalho, principalmente os mos filhos não me dão muito trabalho” (P04.F1B)

Além dos temas abordados até aqui, os entrevistados demonstram percepções sobre as mudanças culturais e sociais em Angola, expressando opiniões sobre o tema:

(77) “E nós antigamente, este mai velho era respeitado. Você chegar ao lado do mai velho não vai cumprimentar, os jovens dagora quando vê mai velho “bom dia”. Nós antigamente não, pra ir cumprimentar o mai velho memo que você não lhe conhece, chega tem que dar a mão, tens que abaixar” (P10.F3B)

Por fim, os tempos trouxeram o medo. Relatos de assaltos e de violência tingem a realidade de insegurança, deixando no ar uma sensação de vulnerabilidade, um desejo de proteção que parece sempre escapar por entre os dedos.

(78) “Aqui a pouco mesmo, até pouco tempo me assaltaram, (...). Chegaram dentro do carro, viram que era nós, e fizeram aquela bandidagem entre eles e tacaram nosso carro. (...) pegaram o nosso carro e pá. (...) Sacaram as coisas da maioria que estavam lá. São violentos” (P16.F2M)

Esses são apenas alguns dos pontos em comum encontrados nos textos. As entrevistas oferecem um panorama rico e diversificado das experiências e percepções da vida em Angola, abordando desafios, costumes e valores de diferentes gerações.

## 6.2 VOZES DE LUANDA: OS DADOS LINGUÍSTICOS EM FOCO

Nesta seção, são apresentadas as análises e interpretações linguísticas dos dados de fala do *corpus* constituído e os resultados obtidos por meio do programa estatístico *Goldvarb X*.

Após o levantamento das ocorrências de estratégias de indeterminação no *corpus* pesquisado, totalizaram-se 1000 dados, os quais foram submetidos para a primeira rodada no programa *Goldvarb X*, cujos resultados são apresentados na Tabela 01, que segue:

**Tabela 1** - Distribuição geral de cada estratégia de indeterminação do sujeito encontradas no *corpus*

VARIANTES	Ocorrências	Porcentagem
Eu	1	0,1%
Tu	14	1,4%
Você	75	7,5%
Eles	11	1,1%
Nós	181	18,1%
A gente	47	4,7%
Ø+V1PP <sup>66</sup>	270	27,0%
Ø+V2PS	19	1,9%
Ø+V3PS	106	10,6%
Ø+V3PP	104	10,4%
Ø+V+Se	69	6,9%
Ø+VInf.	38	3,8%
VPSA	4	0,4%
SFO <sup>67</sup>	2	0,2%
Formas nominais	59	5,9%
<b>TOTAL</b>	<b>1000</b>	<b>100%</b>

Fonte: Elaboração própria.

Diante dos primeiros resultados, foi possível perceber que na comunidade de Luanda não houve apenas o uso das formas mais tradicionais prescritas pelas GNs, como também das

<sup>66</sup> Ø+V1PP – para essa rodada geral, a variante destacada em questão corresponde apenas ao sujeito sem estar lexicalmente produzido.

<sup>67</sup> SFO: Sujeito na Forma Oracional, a exemplo de “as festas daqui **o é preciso** pessoas convidado ou às vezes aquelas festas pessoas pra entrar” (P01.M1B)

formas mais inovadoras. As estratégias que tiveram maiores frequências foram a “ØV1PP”, com 27%, “nós”, com 18,1%, seguidas de “ØV3PS” e “ØV3PP”, com frequências de 10,6% e 10,4%, respectivamente.

Conforme Duarte (1995), em estudo sobre as formas pronominais de sujeito preenchidas e nulas em dados de fala carioca em Amostras NURC (1992), os resultados apontavam uma expressiva queda na ocorrência de sujeito nulo com um indicativo de mudança em curso com as 2ª e 1ª (a 1ª com os mesmos resultados tanto do singular quanto do plural) pessoas liderando à frente do processo em relação à 3ª pessoa. Vinte e três anos depois, Duarte e Reis (2018), revisitando o mesmo fenômeno agora em dados do PB culto do Rio de Janeiro (Copacabana e Nova Iguaçu), com entrevistas gravadas entre 2009 e 2010 do Projeto Concordância, encontraram resultados que revelaram a segunda pessoa com um quadro de mudança praticamente concluído, cerca de 90% de preenchimento do sujeito nas três faixas etárias selecionadas; e as primeiras pessoas seguindo a mesma direção com índices de 84% para os mais jovens e 77% e 78% para as faixas 2 e 3 respectivamente. Esses resultados mostraram que no PB culto existe um processo de mudança em curso na direção do preenchimento do sujeito referencial definido.

No entanto, os dados de fala do português luandense mostraram resultados contrários aos encontrados por Duarte (1995) e Duarte e Reis (2018), ou seja, houve uma preferência pela forma de sujeito não preenchido, “ØV1PP” (a forma mais conservadora) 27%, contra 18% da forma preenchida “nós”, uma diferença de 9% pontos percentuais. Pode-se argumentar que não é uma diferença expressiva, mas, considerando o fato de que as ocorrências foram encontradas em todas as faixas etárias, com um percentual maior entre os mais jovens, isso pode revelar um provável indicativo de mudança aparente com favorecimento da forma de sujeito não realizado contra a forma de sujeito realizado, mesmo que ainda apresente uma preferência discreta. Há de se destacar também que a presença das formas mais conservadoras do português europeu em diversas construções sintáticas ainda é muito presente nos dados de falantes dos luandenses da capital, o que certamente contribui para a manutenção da forma conservadora quanto à realização do sujeito pronominal.

As estratégias utilizadas para fazer referência às primeiras pessoas do plural, “Ø+V1PP” e “Nós”, totalizaram 45,1% dos dados, conforme exemplos a seguir:

(79) “Ø+V1PP – [...] por exemplo, Ø conseguimos entender [...].” (P15. M2M)

(80) “Nós” – [...] não deu conta que que que nós távamos já cercados com os homens da UNITA[...] (P17.F2M)

enquanto as terceiras pessoas do singular e plural, “ $\emptyset$ +V3PS” e “ $\emptyset$ +V3PP”, somaram 30%.

(81) “ $\emptyset$ +V3PS” – [...] – [...] “é...  $\emptyset$  só **tem** três anos.” (P25.F1S)

(82) “ $\emptyset$ +V3PP” – “ $\emptyset$  **Indicaram** a mim pra ser professora das mães [...]. (P21.F3S)

As formas mais inovadoras como “você” e “a gente”, por exemplo, apresentaram 7,5% e 4,7%, respectivamente, pelos falantes de Luanda.

(83) “[...] antigamente, pra **você** namorar com um homem era só namorar assim, bocalmente.” (P09.F3B)

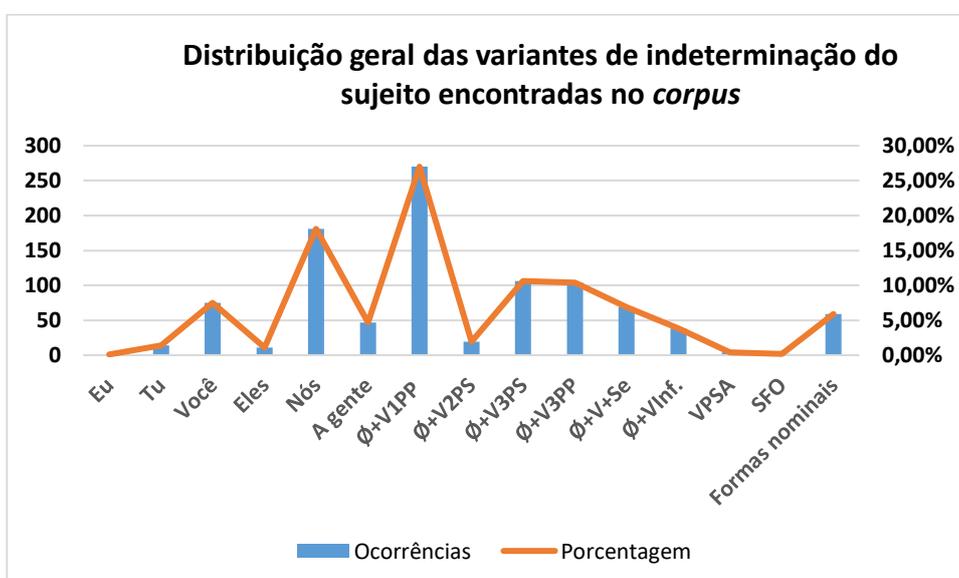
(84) “[...] **a gente** lembra isso que a minha filha canta melhor do que eu.” (P27.M2S)

A forma canônica, a exemplo do “ $\emptyset$ +V+Se”, não se mostrou muito produtiva na amostra. Segue exemplo dessa estratégia:

(85) “[...] **Fala-se** o Português, mas [...] (P15. M2M)

(86) “[...] lá no mato como é, é assim... **constrói-se** assim, a casa grande [do quarto] e a cozinha fica fora [...]. (P17.F2M)

**Gráfico 1** - Distribuição geral das 15 variantes



Fonte: Elaboração própria.

Assim, sendo o fenômeno da indeterminação do sujeito um objeto de estudo nos espaços escolares, os resultados encontrados confirmam a hipótese de que os falantes luandenses também inovam em relação aos usos tradicionais presentes nos compêndios gramaticais, de modo que os mais velhos e os mais jovens fazem o uso de variadas estratégias, condicionadas por fatores sociais e linguísticos

Embora tenha sido encontrada uma grande quantidade de variantes de estratégias para indeterminação do sujeito, optou-se por retirar aquelas cujas ocorrências foram menores, a fim de fazer uma rodada apenas com aqueles com um número maior de ocorrências. Vale destacar que o programa utilizado suporta até nove variantes, assim, outras rodadas foram feitas, considerando uma variante em relação às demais.

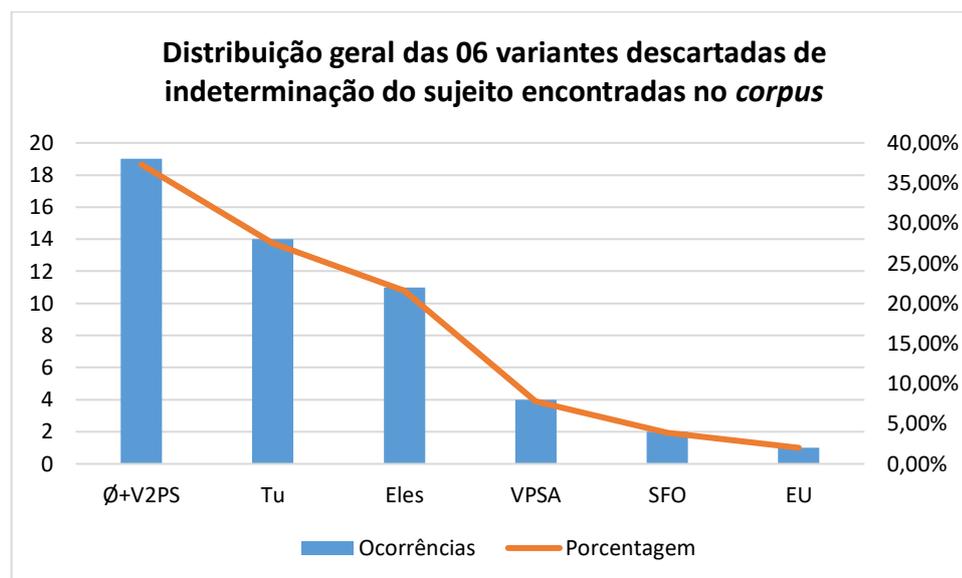
Antes de iniciar as discussões sobre as variantes com maiores frequências, a seguir, apresenta-se uma breve visão da rodada apenas considerando as seis variantes linguísticas descartadas. Ao todo, 51 ocorrências foram encontradas, e, na tabela que segue, podem-se observar as seguintes distribuições e frequências:

**Tabela 2** - Distribuição geral das 06 variantes descartadas de indeterminação do sujeito encontradas no *corpus*

<b>VARIANTES</b>	<b>Ocorrências</b>	<b>Porcentagem</b>
Ø+V2PS	19	37,3%
Tu	14	27,5%
Eles	11	21,6%
VPSA	04	7,8%
SFO	02	3,9%
EU	01	2%
<b>TOTAL</b>	51	100%

Fonte: Elaboração própria.

O gráfico 2, a seguir, permite uma melhor visualização das estratégias descartadas por meio de uma curva decrescente, das mais produtivas às menos produtivas.

**Gráfico 2 - Distribuição geral das 06 estratégias descartadas**

Fonte: Elaboração própria.

Ainda sobre as 06 variantes descartadas, destaca-se que:

1. No grupo 2 – “sexo” – não houve ocorrências da variante “eu” utilizadas por mulheres;
2. No grupo 3 – “faixa etária” – apenas na faixa etária dos mais jovens foram encontradas as formas acima apresentadas, com prevalência das variantes “tu” (12 dados), “Eles” (10 dados), Ø+V2PS (11 dados);
3. No grupo 4 – “escolaridade” – os dados foram prevalentes para o grupo do “Ensino fundamental”, com destaque para a variante “Eles” (09 dados), Ø+V2PS (07 dados) e “tu” (05 dados); não houve ocorrências da variante “Eu” por falantes pertencentes ao Ensino Médio e ao Ensino Superior;
4. No grupo 5 – “Língua de intercomunicação” – os dois fatores apresentaram *knockout*<sup>68</sup> com dados de “eu” para os falantes que falam/compreendem português e falam/compreendem línguas bantu e SFO por falantes apenas da língua portuguesa. No total, foram 37 ocorrências, sendo mais utilizadas por falantes bilíngues com predominância das variantes Ø+V2PS (15 dados) e “tu” e “eles” (05 dados).
5. No grupo 6 – “Grau de Indeterminação” – só não houve *knockout* para a variante “indeterminação parcial/referência implícita”, com 18 e 10 dados de “Ø+V2PS” e “tu”, respectivamente. Assim, não houve ocorrências de “tu”, “eles”, “Ø+V2PS” e “SFO” com “indeterminação completa” e com “indeterminação parcial/referência explícita”.

<sup>68</sup> O termo *Knockout* se refere a “um fator que, num dado momento da análise corresponde a uma frequência de 0% ou 100% para um dos valores da variável dependente” (Guy; Zilles, p. 158, 2007)

6. Nos grupos 7, 8, 10, 11, 13, 14, 16 e 17 – “função da indeterminação”, “preenchimento do sujeito”, “tipo de verbo”, “transitividade verbal”, “tipo de oração”, “tempo e modo verbal”, “mudança e manutenção do referente” e “forma antecedente” – todas as variantes apresentaram *knockout*, sendo os grupos não relevantes para indeterminar o sujeito nas variantes descartadas;
7. No grupo 9 – “Tipo de frase” – não houve ocorrências de indeterminação em frases “Interrogativas” e *knockout* para as variantes “eles”, “eu”, “VPSA” e “SFO” em frases “Negativas”;
8. No grupo 12 – “Núcleo do predicado” – a variante “composto” apresentou também *knockout* diante de “SFO” e “VPSA”. Quando o núcleo do predicado era simples, as maiores ocorrências se apresentaram para as estratégias “ $\emptyset$ +V2PS” (16 dados) e “eles” (10 dados) e “tu” (09 dados) de 41 dados, total.
9. No grupo 15 – “Formas nominais do verbo” – as 05 ocorrências pertenciam ao fator “infinitivo”, configurando, assim, um “*Singleton Group*”<sup>69</sup>.
10. No grupo 18 – “Inclusão do falante” – o fator “Não” inclusão do falante sofreu *Knockout* para as estratégias “SFO” e “eu” e teve sua maior ocorrência diante da estratégia “eles” com 08 dados de um total de 14. Por outro lado, o fator “Sim”, que estabelece a inclusão do falante na indeterminação do sujeito, teve 37 dados, cujas ocorrências foram mais prevalentes nas estratégias “ $\emptyset$ +V2PS” (17 dados), e “tu” (12 dados), restando 11 dados distribuídos em “eles” (03 dados), “SFO” e “VPSA”, com 02 dados cada e apenas 01 dado diante da forma “eu”.

Feitos esses esclarecimentos de cunho metodológico, cabe analisar os resultados encontrados quando consideradas as nove variantes restantes. Assim, de um total de 1000 dados encontrados na amostra, no que concerne às estratégias de indeterminação do sujeito utilizadas pelos residentes de Luanda, foi feito o descarte das 06 variantes (51 ocorrências) que não se mostraram tão frequentes. Seus fatores não estarão mais presentes nas próximas rodadas e análises.

As rodadas que seguem foram realizadas com base nas nove variantes linguísticas, totalizando 949 dados, e suas frequências encontradas são apresentadas na tabela, a seguir:

---

<sup>69</sup> *Singleton Group* refere-se a um grupo quando existem apenas ocorrências para um só fator de uma variável.

**Tabela 3** - Distribuição geral das 09 variantes de indeterminação do sujeito

VARIANTES	Ocorrências	Porcentagem
A gente	47	5%
Nós	181	19,1%
Você	75	7,9%
Ø+V+SE	69	7,3%
Ø+V1PP	270	28,5%
Ø+V3PS	106	11,2%
Ø+V3PP	104	11%
Ø+VINF	38	4,0%
Formas Nominais	59	6,2%
<b>TOTAL</b>	<b>949</b>	<b>100%</b>

Fonte: Elaboração própria.

Observando os dados da rodada com as variantes mais frequentes para o estudo da indeterminação do sujeito, algumas alterações precisaram ser realizadas para a resolução dos problemas de *knockouts* apontados pelo programa em alguns grupos de fatores, no arquivo de condições, a saber:

- (a) Retirou-se do grupo 6 – grau de indeterminação - o fator “indeterminação parcial/referência explícita”, haja vista que só havia 08 dados, 03 de “formas nominais” e 05 de “a gente”;
- (b) No grupo 7 – “função da indeterminação”, ocorreram três *knockouts*. Os fatores “descomprometimento do sujeito”, “ocultação do sujeito” e “economia linguística” não foram representativos na amostra e totalizando apenas 7 dados, dentre eles, com os valores 01, 05 e 01 dados, respectivamente, por isso, foram descartados das rodadas;
- (c) No grupo 8 – “preenchimento do sujeito” os dois fatores “sujeito realizado” e “sujeito não realizado” apresentaram *knockouts*, assim, não houve dados de sujeito realizado para o “a gente” e nem para a forma verbal de infinitivo (ØVINF);
- (d) No grupo 10 – “Tipo de verbo” – Houve *knockout* na variante que corresponde ao verbo “*dicendi*” ligada à estratégia de “ØVINF”. Os dados mais representativos foram os verbos de “ação” e “estado”, com 686 e 177 ocorrências, respectivamente;
- (e) No grupo 11 – “Transitividade verbal” – a única variante que não apresentou *knockout* foi a “transitiva”, com 889 ocorrências. Não houve verbo intransitivo com a

estratégia “nós” e “Ø+V+SE”, bem como verbo de “ligação” com “ØV1PP” “ØV3PS” e “Ø+V+SE”. Sendo assim, o grupo foi descartado;

(f) No grupo 13 – “Tipo de oração” – o *knockout* ocorreu com a oração absoluta em relação à estratégia “ØVINFINF”;

(g) No grupo 14 – “Tempo e modo verbal” – houve *knockouts* para o “pretérito perfeito”, “pretérito imperfeito do indicativo”, “futuro do subjuntivo”, “futuro do presente”, “presente do subjuntivo” e “pretérito imperfeito do subjuntivo”. O presente do indicativo apresentou o maior número de ocorrências na amostra, com 622 dados dos 898. Não houve dados de “imperativo” e nem de “futuro do pretérito”;

(h) No grupo 15 – “Formas nominais do verbo” – houve *knockouts* para as 3 variantes analisadas, sendo que o “particípio” teve apenas 01 dado em relação à estratégia “ØV1PP”, o “gerúndio” com 08 ocorrências distribuídas em apenas 5 variantes. Em relação ao “infinitivo”, apenas as “formas nominais” não apresentaram ocorrências;

(i) No grupo 17 – “Forma antecedente/ paralelismo” – por ser uma variável com 22 variantes, também foi a que mais apresentou *knockouts*. Apenas as variantes “forma precedida por nós”, “forma precedida por ØV1PP” e “forma precedida por um FN explícito” apresentaram valores mais altos de ocorrências, principalmente, diante da flexão verbal -mos (ØV1PP), com 210 dados, também antecedido pela mesma variante “ØV1PP”. Neste grupo, não houve casos de formas precedidas para as seguintes variantes: de “eu explícito” e “eu implícito”, de “forma nominal implícita”, “voz passiva analítica”, “voz passiva sintética” e “Ø+V2PS”;

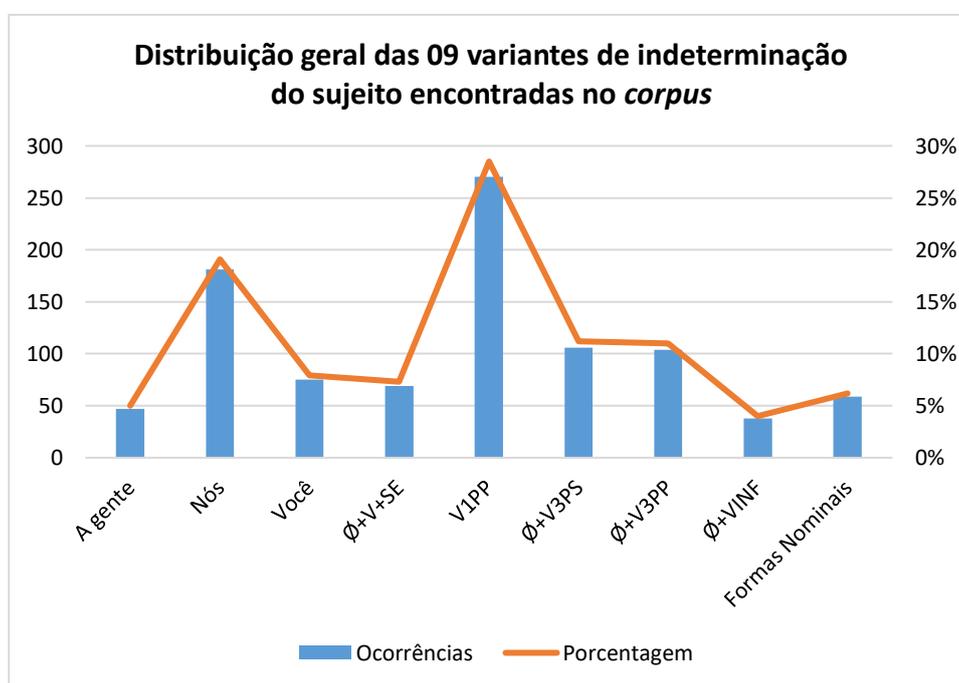
(j) No grupo 18 – “Inclusão do falante” - os resultados apontaram *knockouts* para a “não” inclusão do falante para a variante “nós” e “a gente”. Os dados apontaram 764 ocorrências de “sim”, principalmente diante de desinências -mos “ØV1PP” e “nós”;

(k) No grupo 19 – “Ausência e presença de preposição” - as 64 ocorrências pertenciam ao fator “Presença SE”, configurando, assim, um “*Singleton Group*”. Do total de ocorrências, 56 deles foram encontrados com a variante “Ø+V+SE” e as outras 07 com a variante “V3PS”;

(l) Nos grupos 20, 21 e 22 – “flexão do verbo SE”, “posição do argumento interno do verbo” e “concordância com o argumento interno dos verbos” - os resultados apontaram *knockouts* em todas as variantes, sendo que para todos os 3 grupos que estão relacionados com o SE, como esperado, as maiores ocorrências aconteceram diante das estratégias “Ø+V+SE” e “ØV3PS”, sobretudo, da primeira.

O gráfico, a seguir, apresenta as estratégias que mais foram realizadas pelos entrevistados.

**Gráfico 3** - Distribuição geral das 09 estratégias indeterminadoras



Fonte: Elaboração própria.

Na análise realizada, observou-se que a estratégia “Ø+V+SE”, com 69 ocorrências, foi utilizada pelas 03 faixas etárias, em especial, pelos mais jovens. As frequências encontradas para essas faixas foram: faixa 1: 9,8%, faixa 2: 3,9% e faixa 3: 5,8%. Os falantes da faixa 3 reduzem um pouco o seu uso, mas, mesmo assim, há uma variação estável, pois independentemente da idade de seus falantes, há utilização dessa forma conservadora. Considerando as outras estratégias que foram descartadas, conforme já detalhado, e a distribuição de valores nessa rodada para cada variante, a estratégia “Ø+V+SE”, de acordo o percentual encontrado, os homens representaram 6,7% da amostra, enquanto as mulheres 8,1%. Vale destacar que os participantes que falam e falam/compreendem uma língua bantu apresentaram valores mais baixos de uso para “Ø+V+SE” para indeterminar o sujeito, apresentando um percentual de 5,5%, ao passo que, os falantes que falam apenas a língua portuguesa tenderam a maior utilização dessa forma padrão.

Ainda sobre a variante “Ø+V+SE”, percebeu-se o seu uso nos três níveis de escolaridade. Os menos escolarizados utilizaram bem menos essa estratégia tradicional e, seus percentuais foram crescendo à medida que o nível de escolaridade aumentava, sendo assim, mais produtiva entre os falantes mais escolarizados, o que pode explicar a influência da escolarização ao preservar as formas mais tradicionais no ambiente escolar. Os valores encontrados para o ensino fundamental, ensino médio e ensino superior foram 1,8%, 9,0% e 9,8%, respectivamente.

Uma outra estratégia considerada padrão utilizada pelos participantes luandenses, a terceira pessoa do plural (ØV3PP), foi responsável por 11% do total dos dados analisados. Uma curva decrescente do seu uso aponta que os menos escolarizados usam mais a forma verbal flexionada sem a presença do sujeito, com 18,7%, enquanto os níveis de escolaridade médio e superior fizeram menos o uso dessa forma, apresentando os valores de 9,4% e 6,6%, respectivamente. Para o uso dessa variante, as mulheres apresentaram um valor levemente superior aos homens. Enquanto elas tiveram 11,3%, os homens, 10,7%. Em relação às faixas etárias, o resultado é um pouco parecido com o comportamento das idades com “Ø+V+SE”, embora com valores mais próximos. A faixa 1 foi responsável por 10% do uso e as demais por 13,5% e 10,5%, na ordem. Em relação à língua de intercomunicação, o uso do ØV3PP foi favorecedor entre os falantes que conhecem/falam uma outra língua além da portuguesa.

Sobre algumas das estratégias consideradas não-padrão pelas gramáticas tradicionais, a exemplo do “você” e do “a gente”, a primeira foi mais utilizada pelas mulheres, enquanto, a segunda foi mais utilizada pelos homens. Os percentuais encontrados para essas duas variantes em relação ao sexo feminino foram de 4,2% e de 5,5% para o masculino. Essas duas estratégias inovadoras obtiveram maiores frequências com os falantes menos escolarizados e entre os mais jovens. Em relação ao grupo da escolarização, a variante “a gente” foi reduzida à medida que o participante permaneceu mais tempo na escola. Para os três grupos analisados, o nível fundamental apresentou 11,4%, o médio, 5,7% e o superior, com apenas 0,8%, um valor bem abaixo comparado aos dois primeiros. Em relação à faixa etária, os valores para os três fatores analisados foram 5%, 5,2% e 4,7% para as faixas 1, 2 e 3, respectivamente. Vale destacar que o “a gente” também teve mais frequência diante dos falantes bilíngues do que dos falantes de língua portuguesa, apenas, equivalendo a 6,5% para os primeiros e 1,1% para o grupo “monolíngue”.

Analisando o comportamento da estratégia pronominal “você”, vê-se que os resultados se assemelham com os encontrados do “a gente”. Nessa perspectiva, essa forma mais inovadora também é mais utilizada pelos falantes que compreendem, além do português, uma língua



<i>Concordância com o argumento interno do verbo (SE)</i>								
<i>Sexo</i>							7º	
<i>Faixa Etária</i>		1º				3º	6º	
<i>Escolaridade</i>							5º	
<i>Língua de Intercomunicação</i>						6º		3º

Fonte: Elaboração própria.

Conforme se vê na tabela anterior, alguns grupos de fatores foram mais selecionados do que outros. Os grupos que apresentaram maior recorrência foram os do grupo linguístico: o “tempo e modo verbal”, sendo selecionado pelo “a gente”, “ $\emptyset$ +V3PS”, “ $\emptyset$ +V+SE”, “ $\emptyset$ +VINF” e “formas nominais”; as “formas nominais do verbo”, “ $\emptyset$ +V3PS”, “ $\emptyset$ VINF” e “ $\emptyset$ +V+SE”; a “mudança e manutenção do referente” por “você”, “ $\emptyset$ +V3PS” e formas nominais (para esse último, todos foram a segunda variável selecionada); a forma antecedente/paralelismo teve mais relevância para as variantes e “você” e “ $\emptyset$ +V3PP”; o “tipo de frase”, selecionado pelo “ $\emptyset$ +V3PS”, “ $\emptyset$ +V3PP” e “ $\emptyset$ +V+SE”, na seguinte ordem; o tipo de oração, teve seu destaque sendo selecionado pelo “ $\emptyset$ V3PP”, “ $\emptyset$ +VINF” e “a gente”; Houve, ainda, seleção para o tipo de frase pelo “ $\emptyset$ +V3PS”, “ $\emptyset$ V3PP” e “ $\emptyset$ +V+SE”; o núcleo do predicado foi elegido apenas pelo “ $\emptyset$ +VINF”. Para as duas variáveis semânticas (grau de indeterminação e função da indeterminação. A primeira, foi relevante para a variante “a gente”. A segunda, foi relevante para três variantes: “nós”, “você” e “formas nominais”, todas ela na primeira ordem de seleção.

Em relação às variáveis sociais, apenas quatro variantes foram destacadas: o “a gente”, “ $\emptyset$ +V+SE”, “ $\emptyset$ +VINF” e as “formas nominais”. A “faixa etária”, a “escolaridade” e o “sexo” foram apenas selecionadas pelos fatores “a gente”, o “ $\emptyset$ +V+SE” e “ $\emptyset$ VINF”, respectivamente. A língua de intercomunicação foi selecionada apenas pelas variáveis “formas nominais” e “ $\emptyset$ +V+SE”. Em apenas uma das rodadas, a faixa etária foi a primeira selecionada para a variante “a gente”. Os demais grupos de fatores sociais, em geral, foram selecionados nas últimas colocações.

Para finalizar, destaca-se que não houve seleção para os seguintes grupos pela “ $\emptyset$ V1PP”: preenchimento do sujeito, tipo de verbo, inclusão do Falante, ausência/presença preposição (SE), flexão do verbo (SE), posição do argumento interno do verbo (SE) e a concordância com o argumento interno do verbo (SE).

### 6.2.1 Rodada da estratégia de Indeterminação Nós *versus* demais variantes

Com base nas variantes que já foram descartadas e nas análises já realizadas, optou-se por amalgamar as duas estratégias com maior frequência, utilizadas pelos falantes luandenses para indeterminar o sujeito, “Nós” – forma pronominal e “ØV1PP” – forma não pronominal, partindo do ponto que elas fazem referência a um mesmo elemento, diferenciando-se apenas pelo seu preenchimento ou não preenchimento do sujeito. Desse modo, as seguintes rodadas foram realizadas utilizando essa escolha metodológica.

Antes de mostrar os resultados obtidos para as duas variantes, foi realizada, previamente, uma rodada binária entre essas duas formas. Ao fazer uma rodada contrapondo as estratégias “ØV1PP” e “nós”, o valor da regra de aplicação da variável foi considerada a “ØV1PP” e o valor de *log likelihood* foi de -71.837, conforme se vê na tabela, a seguir:

**Tabela 5** - Influência do ØV1PP no uso da estratégia *nós* indeterminador

<b>Preenchimento do sujeito</b>	Ocorrências	%	Peso Relativo
Sujeito realizado	09/182	4.9%	0.021
Sujeito não realizado	261/269	97%	<b>0.931</b>
TOTAL	270/451	59.9%	<i>Input: 0.708</i>
Nível de significância = 0.000			

Fonte: Elaboração própria.

Diante dessa rodada, verificou-se que apenas um grupo foi selecionado em relação a todos os outros, o grupo do preenchimento do sujeito. Como a regra de aplicação foi a forma verbal da primeira pessoa do plural flexionada, já era de se esperar que a variante “sujeito não realizado” fosse mais favorecedora, apresentando um resultado quase categórico.

Uma rodada complementar também foi realizada, e, ao considerar o “nós” como a regra de aplicação, o valor de *log likelihood* foi o mesmo e os seus resultados, praticamente, foram invertidos, com pouquíssimas diferenças entre os seus valores percentuais e pesos relativos, que pode ser observada na tabela 06.

**Tabela 6** - Influência do *nós* no uso da estratégia ØV1PP indeterminador

<b>Preenchimento do sujeito</b>	Ocorrências	%	Peso Relativo
Sujeito realizado	173/182	95.1%	<b>0.979</b>

Sujeito não realizado	8/269	3%	0.069
TOTAL	181/451	40.1%	<i>Input: 0.292</i>
Nível de significância = 0.000			

Fonte: Elaboração própria.

Diante dessas duas últimas rodadas, verificou-se que tanto a estratégia “ØV1PP” e o “nós”, se comportaram de formas bem parecidas, diferenciando-se apenas pela realização ou não do sujeito. As análises que seguem, como mencionado, anteriormente, foram feitas utilizando as variantes que seguem: “nós”, “você”, “a gente”, “Ø+V3PS”, “Ø+V3PP”, “Ø+V+Se” e as “formas nominais”, cujas ocorrências totalizaram 911 dados. A tabela 7, a seguir, apresenta a nova distribuição e seus percentuais.

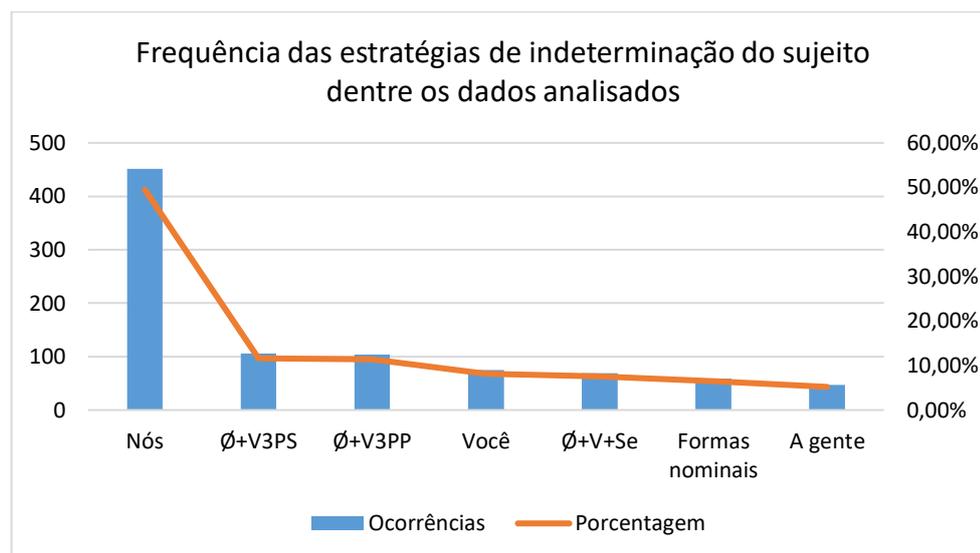
**Tabela 7-** Frequência das estratégias de indeterminação do sujeito dentre os dados analisados

VARIANTES	Ocorrências	Porcentagem
Nós <sup>70</sup>	451	49,5%
Ø+V3PS	106	11,6%
Ø+V3PP	104	11,4%
Você	75	8,2%
Ø+V+Se	69	7,6%
Formas nominais	59	6,5%
A gente	47	5,2%
<b>TOTAL</b>	<b>911</b>	<b>100%</b>

Fonte: Elaboração própria.

O gráfico 04 ilustra as 07 estratégias indeterminadoras com maior produtividade pelos moradores de Luanda.

<sup>70</sup> O termo “nós”, refere-se aos dados da forma pronominal “nós” com a junção dos dados de “ØV1PP” apenas para o grupo da variável dependente. Já, para aos valores que estão ligados às formas antecedentes/paralelismo, há a presença dos fatores “forma precedida de *nós* explícito” e “forma precedida de ØV1PP”, pois, no grupo do paralelismo não houve o almagamento desses dois fatores, por uma escolha metodológica da pesquisadora, cujo objetivo era identificar a real escolha do falante, a forma vazia, como em “ØV1PP”, ou a lexicamente preenchida “nós”, o que pode ser visto em algumas tabelas mais adiante nas seções sobre a influência do paralelismo.

**Gráfico 4** - Distribuição geral das 07 estratégias indeterminadoras

Fonte: Elaboração própria.

Conforme se vê, as primeiras pessoas do plural, o “ØV1PP” e o “nós”, mesmo amalgamadas, os resultados continuam mostrando a grande produtividade dessas formas na comunidade analisada, bem como uma gama de outras estratégias, desde as mais conservadoras e as mais inovadoras.

### 6.2.2 Nós versus as demais

Nesta seção, após feitos os ajustes e o amalgamento das variantes “Nós” pronominal com a “ØV1PP”, tomou-se como regra de aplicação o “nós” para que outras rodadas fossem feitas com as demais, totalizando 451 ocorrências e correspondendo a quase metade das ocorrências encontradas (49,5%).

Assim, nessa primeira rodada, realizou-se a submissão dos dados de “nós” - variante mais canônica - ao programa estatístico *GoldvarbX*, com os dados do “a gente”, uma forma mais inovadora, que, por um processo de gramaticalização, acabou ganhando propriedades mais gramaticais, mais semânticas e funcionais e que, ainda assim, representa uma referência de um eu mais ampliado, mais abrangente em relação aos seus referentes. Após resolução dos *knockouts*, o programa selecionou quatro grupos de fatores por ordem de relevância: escolaridade, função da indeterminação, forma precedida/paralelismo e mudança/manutenção do referente.

### 6.2.3 Nós versus A gente

Para essa rodada, o *log likelihood* foi -86.138, o *input* foi 0.983, com nível de significância 0.041. Vale, inicialmente, destacar que rodada apresentou *knockouts* em relação a alguns grupos de fatores:

1. No grupo “Função da indeterminação” o fator “Economia Linguística”, apresentou apenas 5 dados para a estratégia “nós”;
2. Todos os dados do grupo “Preenchimento do sujeito” ocorreram com o fator “nós”;
3. No grupo do “tempo e modo verbal”, todos os dados do “futuro do presente” e “futuro do subjuntivo”, ocorreram diante de “nós”;
4. Não diferente dos outros grupos já elencados, as “formas nominais do verbo”, apresentaram *knockout* para apenas 01 dado de “gerúndio” e 10 de “infinitivo” para o fator “nós”;
5. O grupo da “forma antecedente/paralelismo” apresentaram *knockouts* por terem pouquíssimas ocorrências nos seguintes fatores: “eles explícito”, “Ø+V3PS”, “nós implícito”, “Ø+V1PP” “você implícito”, “Ø+V2PS” “tu explícito” e “nós implícito”;
6. Para o grupo da inclusão do falante, apenas 01 dado ocorreu quando o “nós” não estava incluso;
7. Os demais grupos “ausência/ presença do se”, “flexão do verbo se”, “posição do argumento interno” e a sua “concordância com o argumento interno se” não apresentaram fatores.

A variável social escolaridade foi a primeira do grupo de fatores selecionada pelo *Goldvarb X*. As variantes para este grupo foram o “ensino fundamental” (nula e baixa), o “ensino médio” e o “ensino superior”. Os dados apontam que essa variável é a que exerce mais influência em relação ao uso do “nós indeterminador” pelos moradores de Luanda, sobretudo para aqueles que permaneceram mais tempo no ambiente escolar e que foram, conseqüentemente, mais expostos às variantes mais normativas. Interessante observar que esse dado pode revelar um indício de que possa estar havendo um processo de mudança, diferentemente de Portugal, que continua a usar mais a estratégia do verbo na terceira pessoa do singular com o se, por exemplo. A hipótese levantada foi a de que o sujeito pronominal “nós” fosse realizado de maneira menos abundante pelos menos escolarizados, tendo como

parâmetro menor contato e menos tempo de exposição ao ensino tradicional e acesso às variantes mais canônicas, ensinadas, majoritariamente, no espaço escolar. Destaca-se que, por um bom tempo na sociolinguística de Angola, nem todos os habitantes tinham acesso direto aos liceus, institutos e faculdades, a menos que estivessem na condição de assimilados.

Seguem exemplos do uso de “nós” nos três tipos de ensino:

#### Ensino Fundamental

(87) “**Nós** tivemos brincadeiras de Agriã, Ø tivemos barra de lenço, [...]” (P05.M2B)

#### Ensino Médio

(88) “**Nós** falamos, por exemplo, conseguimos entender o que o outro fala, [...]”  
(P15. M2M)

#### Ensino Superior

(89) “[...] Então, **nós** estaremos a distribuir esses livros até no período das 12:00 e depois Ø vamos almoçar e só assim que voltaremos a igreja.” (P23.M1S)

A tabela 8, a seguir, mostra a produtividade do “nós” indeterminador em relação ao fator “a gente”:

**Tabela 8** - Influência da escolaridade no uso da estratégia *nós* indeterminador

<b>Escolaridade</b>	Ocorrências	%	Peso Relativo
Ensino Fundamental	86/117	73.5%	0.123
Ensino Médio	151/164	92.1%	0.375
Ensino Superior	214/217	98.6%	<b>0.809</b>
TOTAL	451/498	90.6%	<i>Input</i> : 0.983
Nível de significância = 0.041			

Fonte: Elaboração própria.

Os dados da tabela apontam valores de P.R. crescentes para a estratégia “nós” ao longo do período escolar, favorecido pelo Ensino Superior com 0.809 e sendo desfavorecido pelo Ensino Fundamental com valores um pouco acima de 0,10. Observou-se que, mesmo os estudantes passando mais tempo na escola, a probabilidade de eles utilizarem a forma do “nós” indeterminador foi bem maior, se comparado com aqueles que pouco frequentaram a escola, apresentando assim, menor uso das formas padrão.

A segunda variável selecionada na rodada entre “nós” *versus* “a gente” foi a função da indeterminação. Para essa estratégia, no fator economia linguística, houve *Knockout*, sendo que os cinco casos encontrados foram da estratégia “nós”. Observando os resultados fornecidos pelo programa, a desfocalização do sujeito foi a variante mais produtiva para a indeterminação do sujeito pelos falantes de Luanda, com peso relativo de 0.686, mais alto se comparada as demais formas, que é quando busca focalizar a ação e não o seu sujeito. Em Dias (2017), essa função foi a mais favorecedora em relação à forma pronominal “nós” quando analisou textos escritos baianos dos séculos XIX e XX, elegendo a variável função da indeterminação como a principal favorecedora do seu uso, especialmente quando a intenção é o descomprometimento do agente ou desfocalização do sujeito, centrando, portanto, na ação verbal ou em seu objeto.

**Tabela 9** - Influência da função da indeterminação no uso da estratégia *nós* indeterminador

<b>Função da indeterminação</b>	Ocorrências	%	Peso Relativo
Desfocalização do sujeito	263/279	94.3%	<b>0.686</b>
Exemplificação	171/190	90.0%	0.282
Descomprometimento	2/7	28.6%	0.049
Ocultação do sujeito	10/17	58.8%	0.247
TOTAL	451/498	90.6%	<i>Input: 0.983</i>
Nível de significância = 0.041			

Fonte: Elaboração própria.

#### Desfocalização do sujeito

(90) “Lá no município, nós tirávamos algumas árvores lá no município que tem, como é que se diz” (P17.F2M)

#### Exemplificação

(91) “[...] **nós** na caixa da cozinha sem saber o que fazer [...]” (P17.F2M)

#### Descomprometimento

(92) “[...]. Então cada dia que passa **nós** temos mais alimentos saudáveis [ ...]” (P25.F1S)

#### Ocultação do sujeito

(93) “Foi muito difícil. Nós vivemos quase um ano no *Cuito* e era insuportável” (P29.F2S)

A tabela 10, a seguir, mostra a produtividade do “nós” indeterminador em relação ao fator “a gente”:

A terceira variável selecionada na rodada do “nós” x “a gente” foi a forma precedida/paralelismo que trata sobre a tendência das pessoas repetirem as mesmas estruturas da utilizada anteriormente, já que esse mecanismo propicia uma maior fluência, menor esforço na construção do sentido daquele que produz e da compreensão pelo interlocutor. Os dados a seguir revelam que o uso do “nós” indeterminador teve um maior favorecimento diante da forma precedida de “Ø+V1PP”, com peso relativo de 0.657, enquanto, a forma precedida por “nós explícito”, se aproximou do ponto neutro.

**Tabela 10** - Influência do paralelismo no uso da estratégia *nós* indeterminador

<b>Paralelismo/forma antecedente</b>	Ocorrências	%	Peso Relativo
Forma precedida de <i>nós</i> explícito	91/93	97.8%	0.489
Forma precedida de Ø+V3PP	11/14	78.6%	0.096
Forma precedida de Ø+V1PP	173/175	98.9%	<b>0.657</b>
Forma precedida de <i>a gente</i> explícito	2/9	22.2%	0.011
Forma precedida de <i>a gente</i> implícito	½	50.0%	0.043
Forma precedida de <i>you</i> explícito	¾	75.0%	0.092
Forma precedida de FN explícito	5/6	83.3%	0.211
Forma precedida de Ø+V+Se	5/6	83.3%	0.081
TOTAL	451/498	90.6%	<i>Input: 0.983</i>
Nível de significância = 0.041			

Fonte: Elaboração própria.

Logo a seguir, são mostrados dois excertos que ilustram os contextos mais predominantes na amostra:

Forma precedida de Ø+V1PP

(94) “Praticamente eu digo assim ø já tamos acostumado, **nós** batalhamos memo pra educar os filhos, mas as vezes o Deus dá mesmo muita força” (P04.F1B)

Forma precedida de *nós* explícito

(95) “**Nós** tivemos brincadeiras de Agriã, Ø tivemos barra de lenço” (P05.M2B)

A hipótese aventada para esse grupo era de que uma vez utilizado um pronome no início da cadeia discursiva, esse mesmo pronome condicionaria a outras repetições, independentemente da variante ser uma forma pronominal implícita ou explícita. Os dados

encontrados reforçam essa ideia, visto que as formas “Ø+V1PP” e “nós” referenciam a mesma pessoa, com a diferença que em um, o sujeito está implícito e o outro não. Por exemplo, das 93 ocorrências antecedidas pelo pronome “nós”, 97,8% foram do próprio pronome explícito, havendo resultado parecido para referenciar o “nós” com a forma verbal vazia. Os 175 dados encontrados de “Ø+V1PP” correspondem a quase sua totalidade ao referenciar o “nós”.

O que pode explicar essa diferenciação nos resultados é a codificação realizada, nesse grupo de fatores, pois, ela foi feita distinguindo esses dois fatores. Assim, se fosse feita a junção do “Ø+V1PP” e “nós”, o resultado para a forma precedida “nós”, seria, ainda assim, o maior favorecedor para a manutenção do pronome na cadeia discursiva. As demais formas continuariam com valores inferiores.

Em relação ao “a gente”, as ocorrências implícita/explicita sobre a forma precedida foram ínfimas e seus pesos relativos também.

A quarta variável que se mostrou relevante para a análise do “nós” x “a gente” foi a mudança/manutenção do referente. Essa variável está diretamente ligada com a forma precedida e seus resultados são parecidos, de acordo com o que se vê na tabela 11 a seguir.

**Tabela 11** - Influência da mudança/manutenção do referente no uso da estratégia *nós* indeterminador

<b>Mudança/Manutenção do referente</b>	Ocorrências	%	Peso Relativo
Referente igual à estrutura anterior	157/166	94.6%	<b>0.516</b>
Referente diferente da estrutura anterior	165/174	94.8%	<b>0.762</b>
1ª de uma série	82/98	83.7%	0.232
Forma isolada	47/60	78.3%	0.167
TOTAL	451/498	90.6%	<i>Input: 0.983</i>
Nível de significância = 0.041			

Fonte: Elaboração própria.

Esperava-se que a manutenção do referente acompanhasse a estratégia de indeterminação utilizada pelo falante tivesse um contexto mais favorecedor. Na tabela, percebeu-se que duas variantes foram mais relevantes. O P.R. foi maior diante da utilização das formas diferentes, com valor de 0.762, enquanto, o seu referente era igual, a forma anterior teve um P.R. menor, de 0.516, valor bem próximo ao nível de neutralidade. Como pontuado na seção anterior, esse resultado pode ter sido influenciado pela própria natureza da codificação que diferenciou os diferentes tipos de contextos precedentes, em especial aqueles ligados à primeira

pessoa do plural por meio do uso do “Ø+V1PP” e “nós”. Exemplos são colocados para a ilustração das formas de manutenção ou não da estrutura antecedente:

Referente igual à estrutura anterior

(96) “nunca sabe que **a gente** vai morrer hoje, conforme **a gente** tá. (P05.M2B)

Referente diferente da estrutura anterior

(97) “Ø **Ficamos** sempre com uma pedrinha na mão aquilo **a gente** chamava de mãe né” (P05.M2B)

Quando feita uma análise dos valores percentuais desses dois grupos de manutenção, seus valores são os mais altos do que a “1ª forma de uma série” e de sua “forma isolada”. Destaca-se ainda que essas duas últimas variantes foram as menos favorecedoras visto que a primeira não tem um parâmetro anterior para referência ser igual ou diferente, e a segunda, por ser a única referência naquele contexto discursivo.

#### 6.2.4 Nós versus Você

A rodada “Nós” versus “Você” foi realizada e o programa estatístico selecionou 06 grupos de fatores com a seguinte ordem de relevância: paralelismo, preenchimento do sujeito, do tempo e modo verbal, tipo de verbo, formas nominais do verbo e grau de indeterminação. O nível de significância foi de 0.038, o *log likelihood* foi -83.179, o *input* foi 0.907 para essa rodada. A seguir, destaca-se os *knockouts* apresentados:

1. No grupo “grau da indeterminação” o fator “Indeterminação parcial/referência explícita”, apenas 05 ocorrências para a estratégia “você”;
2. No grupo “função da indeterminação”, o fator “Economia Linguística” e a “Ocultação do sujeito” apresentaram apenas 05 e 10 ocorrências, respectivamente para a estratégia “nós”;
3. No grupo do “tempo e modo verbal”, todos os dados do “futuro do presente” e “pretérito imperfeito do subjuntivo”, ocorreram diante da variante “nós”;
4. O grupo da “forma antecedente/paralelismo” apresentaram *knockouts* por terem pouquíssimas ocorrências nos seguintes fatores: Ø+V3PS, “nós implícito” “a gente implícito”, “Ø+V+SE, “Ø+V2PS” e “tu explícito”

5. Nos demais grupos “ausência/ presença do se”, “flexão do verbo se”, “posição do argumento interno” e a sua “concordância com o argumento interno se” houve *Singleton groups*.

Embora o paralelismo não tenha sido selecionado pelo programa *GoldvarbX* como o grupo que mais contribuiu para a indeterminação do sujeito na rodada anterior entre “nós” *versus* “a gente”, pois ele foi o 3º grupo selecionado dentre os 4). Percebeu-se que, em relação ao “nós” *versus* “você”, a forma antecedente já se comportou de maneira mais expressiva, sendo o grupo que mais influenciou as estratégias de indeterminação nesta rodada. Dentre os nove fatores analisados, as formas precedidas pelo “Ø+V1PP” e “nós” explícito foram as que menos inibiram o fenômeno da indeterminação pelos luandenses, com destaque para o primeiro que alcançou P.R de 0.801, enquanto, o segundo favoreceu, com um valor de P.R. um pouco mais abaixo de 0.565.

**Tabela 12** - Influência do paralelismo no uso da estratégia *nós* indeterminador

<b>Paralelismo/forma antecedente</b>	Ocorrências	%	Peso Relativo
Forma precedida de <i>nós</i> explícito	91/94	96.8%	<b>0.565</b>
Forma precedida de Ø+V3PS	10/13	76.9%	0.090
Forma precedida de Ø+V1PP	173/174	99.4%	<b>0.801</b>
Forma precedida de <i>a gente</i> explícito	2/3	66.7%	0.065
Forma precedida de Ø+VInf.	9/12	75.0%	0.162
Forma precedida de <i>você</i> explícito	3/33	9.1%	0.005
Forma precedida de <i>você</i> implícito	¼	25.0%	0.004
Forma precedida de FN explícito	5/7	71.4%	0.243
Forma precedida de <i>eles</i> explícito	½	50.0%	0.022
TOTAL	295/342	86.3%	<i>Input: 0.907</i>
Nível de significância = 0.038			

Fonte: Elaboração própria.

O preenchimento do sujeito foi o segundo grupo desta rodada que apontou maior relevância. Analisando os dados de Luanda, nota-se que os falantes do português que moram na capital têm utilizado as formas mais nulas, as formas vazias, o que pode colaborar para que em situações contextuais diversas as formas nulas também possam ter uma preferência, em detrimento das formas com sujeito preenchido.

Os dados mostram que o fator “sujeito realizado” é o que menos inibe o fenômeno da indeterminação do sujeito, com 0.132 de peso relativo contra 0.851 em relação a sua forma preenchida.

**Tabela 13** - Influência do preenchimento do sujeito no uso da estratégia *nós* indeterminador

<b>Preenchimento do sujeito</b>	Ocorrências	%	Peso Relativo
Sujeito realizado	182/253	71.9%	0.132
Sujeito não realizado	269/273	98.5%	<b>0.851</b>
TOTAL	451/526	85.7%	<i>Input: 0.907</i>
Nível de significância = 0.038			

Fonte: Elaboração própria.

O terceiro grupo que mais contribuiu para o favorecimento da indeterminação do sujeito pelos falantes de Luanda, em relação ao uso da estratégia “nós”, foi o grupo do tempo e modo verbal. Muito embora o presente do indicativo tenha sido o fator com o maior número de ocorrências nas entrevistas analisadas, o seu peso relativo não foi alto, ficando os tempos pretéritos nas primeiras posições. Importante observar que a estratégia “nós” foi, predominantemente, influenciada pelo modo indicativo. O fator da primeira posição, de acordo com os dados fornecidos pelo programa, o pretérito imperfeito do indicativo, apresentou P.R. de 0.875, enquanto o pretérito perfeito do indicativo vem logo após, com P.R. de valor 0.796. Esse resultado, possivelmente, se deve à própria natureza das entrevistas, nas quais os entrevistados relataram fatos acontecidos e cotidianos desempenhados no atual momento ou nas suas recordações em tempos passados. A seguir, os dados são apresentados na tabela 14.

**Tabela 14** - Influência do tempo e modo verbal no uso da estratégia *nós* indeterminador

<b>Tempo e modo verbal</b>	Ocorrências	%	Peso Relativo
Presente do Indicativo	297/351	84.6%	0.360
Pretérito Imperfeito do Indicativo	84/86	97.7%	<b>0.875</b>
Futuro do Subjuntivo	8/12	66.7%	0.072
Pretérito Perfeito do indicativo	45/48	93.8%	<b>0.796</b>
TOTAL	434/497	87.3%	<i>Input: 0.907</i>
Nível de significância = 0.038			

Fonte: Elaboração própria.

Seguem os exemplos do pretérito imperfeito em (104) e o perfeito em (105).

(98) “à tardinha Ø brincávamos a macaca, ringue, essas brincadeiras [...]” (P07.F2B)

(99) “nós **corremos** o risco de poder contrair estas mesmas doenças” (P11.M1M)

Um outro grupo relevante para a variante “nós” foi o “tipo de verbo” mais utilizado nos discursos dos entrevistados como indeterminador. Dentre os quatro fatores analisados, os verbos de “estado”, juntamente com os verbos “epistêmicos” foram os que apresentaram não somente porcentagens altas para o número de ocorrências, como também foram os que mais favoreceram à regra de aplicação da variante “nós”, os primeiros com peso relativo de 0.868, acompanhados pelo peso relativo de 0.631, do segundo colocado. A seguir, encontram-se excertos dos contextos favorecedores e a tabela com os respectivos valores encontrados para esse grupo.

(100) “[...] mesmo hoje **a gente** tá, [...]” (P02.M1B)

(101) “[...] nós não **conseguimos**, não temos aquela tradução exata da coisa [...]” (P12.M1M)

**Tabela 15** - Influência do tipo de verbo no uso da estratégia *nós* indeterminador

<b>Tipo de verbo</b>	Ocorrências	%	Peso Relativo
Ação	286/346	82.7%	0.309
Estado	137/143	95.8%	<b>0.868</b>
Epistêmicos	24/29	82.8%	<b>0.631</b>
<i>Dicendi</i>	4/6	66.7%	0.253
TOTAL	451/524	86.1%	<i>Input: 0.907</i>
Nível de significância = 0.038			

Fonte: Elaboração própria.

Embora, frequentemente utilizados pelos participantes da pesquisa, e por apresentar, naturalmente, um vasto número de verbos na língua portuguesa, os verbos de “ação”, mesmo ultrapassando a quantidade de ocorrências quando somados aos demais fatores, foram inibidores. Não somente, esperava-se que os verbos “*dicendi*” apresentassem valores também mais altos, visto que muitas das falas e situações vividas por eles, como o resgate da memória ocorria por meio das retextualizações.

O quinto grupo com grande relevância na utilização da variante “nós” como uma estratégia de indeterminação largamente utilizados pelos luandenses foi a forma nominal em

que o verbo se encontrava. Mesmo não sendo tão presente na amostra, a forma infinitiva, apresentou um leve favorecimento de 0.561 de peso relativo em relação aos demais formas. A fator “particípio” não teve ocorrências. Com apenas 0.071 de peso relativo, o gerúndio teve apenas uma ocorrência que pode ser visualizada em:

(102) “[...]ø **metendo** as vigotas por cima e é muito pesado, [...]” (P02.M1B)

O “grau de indeterminação” foi o sexto e último grupo selecionado para a estratégia “nós” x “você”. Esperava-se encontrar na amostra uma maior produção da indeterminação completa, visto que qualquer pessoa poderia ser o sujeito. No entanto, o grau que se mostrou mais produtivo foi quando se havia um tipo de indeterminação parcial e, nesse caso, com o referente implícito. Deste modo, os contextos de indeterminação tiveram maiores favorecimentos quando foi possível identificar por meio de inferência pelo menos um dos seus referentes, obtendo um P.R. pouco acima do ponto neutro, 0.536, com 82.7% dos dados. O contexto em que foi impossível identificar os seus referentes – indeterminação completa – ficou bem próximo à marca de 67%, no entanto, com peso relativo que não o favoreceu.

Destaca-se que, diante das análises já realizadas para as duas primeiras rodadas, “nós” x “a gente” e “nós” x “você”, houve a presença apenas de uma variável social selecionada que foi a escolaridade, em primeira posição para o primeiro grupo. Assim, vê-se que os resultados por ora apresentados podem indicar que o fenômeno da indeterminação do sujeito em Luanda, tem sido influenciada mais por variáveis mais linguísticas do que as extralinguísticas.

### 6.2.5 Nós *versus* Formas nominais

Uma outra rodada foi feita, “nós” e as “formas nominais”, e o programa selecionou oito grupos de fatores na seguinte ordem: inclusão do falante, grau de indeterminação, preenchimento do sujeito, sexo, tempo e modo verbal, tipo de verbo e, por fim, o tipo de oração. O nível de significância foi de 0.037 e o *input* 0.092. Dentre todas as variáveis, mais uma vez, se tem uma predominância de seleção de contextos linguísticos do que sociais. Para essa análise, analisou-se a inclusão do falante como um dos contextos mais favorecedores, com destaque para situações em que o falante luandense se coloca como parte integrante a uma situação mais generalizadora.

Assim, a maioria dos casos em que houve a indeterminação do sujeito, a “não inclusão” do falante ao discurso foi praticamente inibida com 0.001 de peso relativo com menos de 4% de ocorrências

enquanto o contexto em que a referência do sujeito indeterminado foi mais produtivo, com peso relativo de 0.594 e 93.4%. Nos dados, o uso das “formas nominais”, o falante foi raramente incluso no universo de referência.

Na segunda variável selecionada, o grau de indeterminação, o percentual foi maior quando a indeterminação do sujeito era parcial e tinha referência implícita, com mais de 96% dos dados. Para essa variante, o valor do peso relativo também foi alto, sendo o maior favorecedor de “nós”. Os dados podem refletir o fato de que os contextos que indicam o grau de indeterminação parcial estejam diretamente relacionados ao contexto referente do sujeito indeterminador tenha o falante incluído nele.

**Tabela 16** - Influência do grau de indeterminação no uso da estratégia *nós* indeterminador

<b>Grau de indeterminação</b>	Ocorrências	%	Peso Relativo
Indeterminação parcial/referência implícita	415/432	96.1%	<b>0.645</b>
Indeterminação completa	36/75	48.0%	0.031
TOTAL	451/507	89.0%	<i>Input: 0.992</i>
Nível de significância = 0.037			

Fonte: Elaboração própria.

O “preenchimento do sujeito” foi a terceira variável selecionada pelo *GoldvarbX*. Nessa variável, o sujeito indeterminado foi favorecido de forma bastante significativa diante da estratégia não realizada lexicalmente, com P.R de 0.795 e quase 99% dos dados e, para a forma preenchida lexicalmente, o P.R. alcançou apenas 0.175. Vale destacar aqui, que os dados de “Ø+V1PP” foram amalgamados com o “nós”, por isso, essa junção pode explicar um valor mais alto para as formas que não apresentaram o seu referente, não tendo assim o sujeito explícito no discurso.

O “sexo”, nesta rodada, foi a primeira e única variável social selecionada na rodada da estratégia “nós” em relação às “formas nominais”. As ocorrências, a porcentagem e o peso relativo do “nós” indeterminador foi bastante produtivo na fala dos homens. Enquanto os homens utilizaram a estratégia “nós” com mais frequência, obtendo um P.R de 0.655, as mulheres foram as mais inibidoras ao seu uso, com PR. 0.265.

O “tempo e modo verbal” foi a quinta variável selecionada e seus dados são bem parecidos à rodada do “nós” x “você”. Nesse grupo apenas 03 fatores foram selecionados e abarcam os modos infinitivos, com predominância de pesos relativos maiores diante das formas

pretéritas, mesmo elas não tendo as maiores ocorrências como as do tempo presente, como se observa a seguir:

**Tabela 17** - Influência do tempo e modo verbal no uso da estratégia *nós* indeterminador

<b>Tempo e modo verbal</b>	Ocorrências	%	Peso Relativo
Presente do Indicativo	297/345	86.1%	0.394
Pretérito Imperfeito do Indicativo	84/88	95.5%	<b>0.654</b>
Pretérito Perfeito	45/46	97.8%	<b>0.881</b>
TOTAL	426/479	88.9%	<i>Input: 0.992</i>
Nível de significância = 0.037			

Fonte: Elaboração própria.

Destaca-se ainda que, na variável “tipo de verbo”, a sexta selecionada pelo programa, os verbos de “ação”, muito embora tenham sido os mais produtivos, apresentaram valores inibidores, enquanto os verbos epistêmicos e de estado foram os que mais favoreceram, com 0.934 e 0.790, respectivamente. Resultado parecido com a rodada feita com o “você”.

Nessa rodada de “nós” com as “formas nominais” o programa apontou um grande favorecimento das orações “principal” e “absoluta” na indeterminação do “nós” quando se discute a variável “tipo de oração”. As orações coordenadas e subordinadas ficaram um pouco próximo ao valor do ponto neutro, com 0.414 e 0.400 de peso relativo. Como já mencionado, o destaque aconteceu com a variante “oração principal” obtendo uma regra quase categórica, com 0.976 de P.R, sendo seguido por 0.583 de P.R. das orações coordenadas.

**Tabela 18** - Influência do tipo de oração no uso da estratégia *nós* indeterminador

<b>Tipo de oração</b>	Ocorrências	%	Peso Relativo
Absoluta	59/69	85.5%	<b>0.583</b>
Coordenada	309/339	91.2%	0.414
Principal	30/33	90.9%	<b>0.976</b>
Subordinada	53/69	76.8%	0.400
TOTAL	451/510	88.4%	<i>Input: 0.992</i>
Nível de significância = 0.037			

Fonte: Elaboração própria.

Seguem alguns exemplos encontrados na amostra:

(103) “[...] minha filha, **as pessoas** anda já (...) não tem mesmo meios. [...]” - (P09.F3M)

(104) “[...] **Você tem** que suportar porque tá a fazer o barulho[...]

A Influência da mudança/manutenção do referente no uso da estratégia “nós” indeterminador, foi amplamente favorecido quando o seu referente era o mesmo, obtendo percentuais e pesos relativos mais altos, com 0.810, enquanto, quando o “referente era diferente”, era a primeira parte de uma série ou até mesmo quando fazia parte de uma estrutura isolada, houve inibição com os seguintes pesos relativos informados, em ordem: 0.425, 0.298 e 0.188.

### 6.2.6 Nós versus Ø+V3PS

Por meio de uma rodada entre as estratégias “nós” e “Ø+V3PS”, observou-se:

1. No grupo “função da indeterminação” os fatores “Economia Linguística”, a “Ocultação do sujeito” também apresentaram apenas 05 e 10 ocorrências, respectivamente para a estratégia “nós”, com já apontado anteriormente;
2. O grupo da “transitividade verbal”, o fator verbo de ligação sofreu *knockout*.
3. No grupo do “tempo e modo verbal”, todos os dados do “futuro do subjuntivo” e “pretérito imperfeito do subjuntivo”, ocorreram diante da variante “nós”;
4. O grupo da “forma antecedente/paralelismo” apresentaram *knockouts* por terem pouquíssimas ocorrências nos seguintes fatores: “eles explícito”, “Ø+V3PS”, “a gente implícito”, “tu explícito” e “eu”.
5. Nos demais grupos “ausência/ presença do se”, “flexão do verbo se”, “posição do argumento interno” e a sua “concordância com o argumento interno se” houve *Singleton groups* .

A rodada pelo *GoldvarbX* com *log likelihood* -113.672, nível de significância 0.017 e seu *input* 0.875, selecionou 6 grupos nesta ordem: “inclusão do falante”, “preenchimento do sujeito”, “paralelismo/forma antecedente”, “mudança/manutenção do referente”, “função da indeterminação” e as “formas nominais do verbo”.

Para o primeiro grupo selecionado, foi a “inclusão do falante” com P.R. de 0.574 quando o falante se incluía no contexto representando quase 90% dos dados encontrados na amostra, enquanto para a “não inclusão”, o P.R. foi bem baixo, 0.011 e com menos de 3% de ocorrências.

O “preenchimento de sujeito”, o segundo grupo selecionado para esta rodada, como a maioria dos resultados apresentados nas outras aqui já realizadas, teve uma grande influência para a realização do seu sujeito para a variante “nós” com PR 0.962 e 98.9% dos dados. No entanto, ele não foi favorecido quando o sujeito não era dito na cadeia discursiva, alcançando PR de apenas 0.169.

Sobre a terceira variável selecionada, o paralelismo/forma precedente foi bastante relevante quando o “nós” era precedido por ele mesmo em sua forma lexicalmente preenchida e quando era precedida pela forma vazia, também da primeira pessoa do plural “Ø+V1PP”. O valor encontrado para o primeiro obteve um peso de 0.769 e, o segundo, 0.711, conforme se vê na tabela 19, a seguir:

**Tabela 19** - Influência do paralelismo no uso da estratégia *nós* indeterminador

Paralelismo/forma antecedente	Ocorrências	%	Peso Relativo
Forma precedida de <i>nós</i> explícito	91/94	96.8%	<b>0.769</b>
Forma precedida de Ø+V3PS	10/34	29.4%	0.088
Forma precedida de Ø+V3PP	11/13	84.6%	0.467
Forma precedida de Ø+V1PP	173/181	95.6%	<b>0.711</b>
Forma precedida de <i>a gente</i> explícito	2/3	66.7%	0.359
Forma precedida de Ø+VInf.	9/12	75.0%	0.258
Forma precedida de <i>you</i> explícito	3/11	27.3%	0.014
Forma precedida de <i>you</i> implícito	1/4	25.0%	0.048
Forma precedida de FN explícito	5/16	31.2%	0.006
Forma precedida de Ø+V2PS	1/3	33.3%	0.017
Forma precedida de Ø+V+Se	5/12	41.7%	0.086
TOTAL	311/383	81.2%	<i>Input: 0.875</i>
Nível de significância = 0.017			

Fonte: Elaboração própria.

Relacionada ao grupo selecionado anteriormente, a influência da mudança/manutenção do referente no uso da estratégia *nós* indeterminador é favorecido quando se tem formas isoladas (0.793), a primeira de uma série (0.748) ou quando o seu referente é igual à estrutura anterior (0.561) e fortemente desfavorecido quando o seu referente é diferente da estrutura anterior. (0.272).

A quinta variável selecionada foi a “função da indeterminação” no uso da estratégia *nós* indeterminador, sendo favorecido quando havia uma exemplificação de uma situação descrita, obtendo um P.R de 0.655. Para a “desfocalização” do sujeito e o “descomprometimento” os pesos foram 0.430 e 0.047, respectivamente. O último grupo destacado pelo programa foi as “formas nominais” sendo produtivo diante da forma infinitiva (0.530) enquanto o gerúndio alcançou (0.430).

### 6.2.7 Nós versus Ø+V3PP

Para a rodada de dados “nós” x “Ø+V3PP”, o *log likelihood* fornecido pelo programa foi -105.912, o *input*: 0.907 com um nível de significância de 0.001. Para esse confronto entre as duas estratégias, foram selecionadas quatro variáveis linguísticas, a saber: inclusão do falante, preenchimento do sujeito, tempo e modo verbal e tipo de verbo.

A variante “inclusão do falante” tem sido selecionada em grande parte nas rodadas realizadas, já que muitas vezes o falante indetermina o sujeito não de forma completa, mas sim parcialmente, sendo ele de caráter implícito ou explícito na cadeia discursiva. Seguindo resultados parecidos, a forma pronominal “nós” quando confrontada com a terceira pessoa do plural sem estar preenchida lexicalmente, foi também bastante influenciada quando o próprio falante se incluía no contexto, alcançando um PR. de 0.698 e representando 92,8% das ocorrências, ao mesmo tempo que a não inclusão apresentou um P.R muito baixo, 0.003, com 01 dado apenas.

Sobre a variável “Preenchimento do sujeito”, o destaque ficou para o sujeito realizado, visto que, os dados de “nós” naturalmente, já podem estar explícitos na fala das pessoas, enquanto a terceira pessoa do plural não exige essa completude, principalmente, em seu caráter indeterminador. Os dados apresentaram 0.787 de P.R para a forma realizada e 0.341 quando o sujeito não foi realizado.

Em relação aos “tempos e modos verbais”, apenas o “pretérito imperfeito do indicativo” não se mostrou significativo para essa rodada, com 0.205 de peso relativo. Os fatores que mais

influenciaram foram o “futuro do subjuntivo”, seguido pelo “presente do indicativo” e, logo após, o “pretérito perfeito”, com os pesos em ordem: 0.604, 0.596 e 0.532.

Semelhantemente às outras rodadas discutidas anteriormente, os verbos de “estado” e “epistêmico” foram os mais favorecedores para o uso do “nós” na rodada em relação com o “ $\emptyset$ +V3PP” e tiveram pesos de 0.710 e 0.517.

### 6.2.8 Nós *versus* $\emptyset$ +V+Se

Para esta rodada de *log likelihood* -75.036, nível de significância 0.015 e *input* 0.991, entre o “nós” e a forma vazia com a partícula se ( $\emptyset$ +V+Se), apontaram sete variáveis selecionadas.

Na primeira, “inclusão do falante”, os dados se repetem para o “nós” com pesos mais altos quando o sujeito era incluído, enquanto o peso relativo 0.000 aconteceu diante do fator “não inclusão” do sujeito, indicando a ausência de variação. Já, o segundo grupo destacado foi o “paralelismo”, em que as formas precedidas por formas nominais explícitas, diante de “VINP” e, por fim da estratégia “ $\emptyset$ +V1PP”. Vale destacar que os valores dos pesos relativos encontrados apresentaram valores bem mais altos que as demais rodadas obtendo-se 1.000, 0.999 e 0.694, na ordem apresentada, bem diferente das formas “a gente explícito” e do “você explícito” com pesos 0.001, para ambos. Esses valores encontrados apontam para a não variação da variante em questão.

No terceiro grupo, o “sujeito realizado” continuou sendo bastante favorecedor da estratégia “nós” indeterminadora sobre o “preenchimento do sujeito”, com peso relativo de 0.838 e, 0.283 para o “sujeito não realizado”.

O quarto grupo destacado foi o “tipo de oração”, sendo que apenas a oração principal alcançou o peso relativo máximo (1.000), sendo a mais produtiva no uso da indeterminação pela estratégia “nós”, embora não tenha tido o maior número de ocorrências na amostra. Situação parecida foi encontrada nos dados que correspondiam ao ensino fundamental, que ficou bem próximo ao peso relativo máximo (0.991) em relação aos fatores do ensino médio e superior. Para esses últimos, os pesos relativos equivaleram a 0.313 e 0.243, respectivamente.

O penúltimo grupo selecionado nesta rodada do “tipo de verbo”, os de estado, favoreceram mais do que os epistêmicos, com 0.788 e 0.554. Na última posição, houve a seleção de uma função semântica - o grau de indeterminação - sendo favorecidos pela variante

indeterminação parcial/referência implícita, com 0.545 e 0.186 para os contextos em que havia a indeterminação completa.

### 6.2.9 Nós versus Ø+VINF

A última rodada foi feita contrastando a variante “nós” com a forma vazia do “infinitivo”. O *log likelihood* estabelecido pelo programa estatístico para essa rodada foi de -32.304, o *input* foi 0.493 e 0.012 para o nível de significância. Para essa rodada, seis variáveis foram selecionadas, sendo que as quatro primeiras foram linguísticas (inclusão do falante, paralelismo/forma antecedente, formas nominais do verbo e mudança ou manutenção do referente e as duas restantes foram extralinguísticas (a língua de intercomunicação e a escolaridade).

Para o primeiro grupo linguístico selecionado está a inclusão do falante. Os resultados encontrados convergem com outras análises já realizadas em que muitos casos o falante, mesmo em uma situação indeterminadora, se inclui no fato realizado. O falante incluso apresentou 0.549 de P.R. enquanto, ele não se incluía inibiu a regra com apenas 0.001 de PR, valor praticamente categórico.

O segundo grupo, o paralelismo/forma precedida, foi a segunda estratégia selecionada. Para essa variável os dados apresentaram grande favorecimento do “nós” diante de uma forma precedida por uma forma nominal, com 0.924 de peso relativo, sendo seguido por uma outra forma precedida, pelo “nós explícito”, como também de uma forma precedida vazia de referência à primeira pessoa do plural (Ø+V1PP), com P.R. de 0.669 e 0.649, respectivamente, conforme evidenciado na tabela 20, a seguir:

**Tabela 20** - Influência do paralelismo no uso da estratégia *nós* indeterminador

<b>Paralelismo/forma antecedente</b>	Ocorrências	%	Peso Relativo
Forma precedida de <i>nós</i> explícito	91/93	97.8%	<b>0.669</b>
Forma precedida de <i>a gente</i> explícito	2/4	50%	0.042
Forma precedida de <i>tu</i> explícito	½	50%	0.023
Forma precedida de <i>você</i> explícito	3/7	42.9%	0.043

Forma precedida de <i>você</i> implícito	½	50%	0.064
Forma precedida de <i>eles</i> explícito	½	50%	0.338
Forma precedida de FN explícito	5/6	83.3%	<b>0.924</b>
Forma precedida de Ø+V3PS	10/15	66.7%	0.048
Forma precedida de Ø+V1PP	173/176	98.3%	<b>0.649</b>
Forma precedida de Ø+V2PS	¼	25%	0.005
Forma precedida de Ø+VInf.	9/15	60%	0.018
Forma precedida de Ø+V+Se	5/8	62.5%	0.138
TOTAL	302/334	90.4%	<i>Input: 0.493</i>
Nível de significância = 0.012			

Fonte: Elaboração própria.

O terceiro grupo de relevância para essa rodada foram as formas nominais do verbo, com ênfase para a variante gerúndio, marcando 0.976 de peso relativo, ao passo que o infinitivo alcançou 0.441 de peso relativo, sendo assim, uma variante mais inibidora.

O último grupo que se mostrou importante foi a “mudança ou a manutenção do seu referente” na rodada “nós” e “Ø+VINFIN.”. Com valores bem altos, com peso relativo 0.973 e 0.910, respectivamente, o uso do “nós” foi favorecido diante de uma forma isolada e quando a estrutura se repetia ao longo da cadeia discursiva, marcando assim manutenção do referente igual à estrutura anterior.

O quinto grupo de fatores selecionado foi também o primeiro grupo de variável social para essa rodada, a língua de intercomunicação. Sobre ela, percebeu-se que os moradores de Luanda, quando falam apenas a língua portuguesa, o uso da estratégia “nós” é menos favorecido, tendo um peso relativo de apenas 0.173, ao contrário daqueles que falam ou compreendem uma língua nacional, sendo estes os maiores favorecedores da regra de aplicação do “nós” com um peso relativo de 0.713.

**Tabela 21** - Influência da língua de intercomunicação no uso da estratégia *nós* indeterminador

<b>Língua de intercomunicação</b>	Ocorrências	%	Peso Relativo
Fala apenas português	164/180	91.1%	0.173
Fala português e compreende/fala língua(s) bantu	287/309	92.9%	<b>0.713</b>
TOTAL	451/489	92.2%	<i>Input: 0.493</i>
Nível de significância = 0.012			

Fonte: Elaboração própria.

Por fim, a última variável selecionada diz respeito ao nível de escolaridade ao qual os indivíduos se encontram. Os dados aqui apontam o favorecimento do “nós” para os grupos que já passaram mais tempo no ambiente escolar, como é o caso dos níveis médio e superior. Esses dois fatores apresentaram pesos relativos de 0.669 e 0.578, respectivamente, enquanto os menos escolarizados inibiram a aplicação da regra, com um PR. bem abaixo dos demais, com 0.108, conforme se vê na tabela 22, a seguir.

**Tabela 22** - Influência da escolaridade no uso da estratégia *nós* indeterminador

<b>Escolaridade</b>	Ocorrências	%	Peso Relativo
Ensino Fundamental	86/98	87.8%	0.108
Ensino Médio	151/158	95.6%	<b>0.699</b>
Ensino Superior	214/233	91.8%	<b>0.578</b>
TOTAL	451/489	92.2%	<i>Input: 0.493</i>
Nível de significância = 0.012			

Fonte: Elaboração própria.

### 6.3 FORMAS NOMINAIS

Conforme explicitado na seção anterior, as formas nominais são estratégias que estão aparecendo cada vez mais na fala das pessoas como um meio de omitir o nome da pessoa, indeterminando assim o referente do seu sujeito. Dias (2011, p. 04) destaca que a variante FN não são classificadas “como pronominais, pois não exercem essa função, tão pouco serem comparadas com outras estratégias não-pronominais, formadas principalmente por um verbo sem sujeito lexical expresso, sendo que elas ocupam a função de sujeito preenchido”. Nesta pesquisa foram encontradas diferentes tipos de sintagmas nominais como o “a(s) pessoa(s)”, “o pessoal”, “o sujeito”, “o povo” como já encontrados nos dados de Carvalho (2010), Assunção (2012) e Souza (2014), entre outros, além de termos como “a criança”, “homem” e “fulana”, com caráter mais generalizador.

Sobre as formas “a(s) pessoa(s)” e “o pessoal”, com destaque à primeira,

#### **A(s) pessoa(s), o pessoal**

(105) “mas antigamente nós tínhamos o divertimento que era a discoteca, **a pessoa** divertia um bocado” (P03.F1B)

(106) “Epa por mim tá muito caro né, mas eles também não são culpados, não são culpado porque, porque as vez **a pessoa** quando diz assim no taxista, ah, mas poxa, o táxi está muito caro ele diz assim ‘não o, o problema e a gasolina, como é?’” (P04.F1B)

(107) “Só é difícil quando **a pessoa** não trabalha. Digo assim né. Diz se a pessoa tiver trabalhar. (P05.M2B)

(108) “Os taxistas, quer dizer... minha filha, **as pessoas** anda já, não tem mesmo meios. Nós somos pobres, as outras têm, nós não temos.” (P09.F3B)

(109) “**A pessoa** que vai conhecer mesmo, vai namorar dois anos ou um ano, nada, nada, nada mesmo.” (P10.F3B)

(110) “levam os passageiros de uma forma imprópria, é... daí **o pessoal** manda parar o táxi” (P21.F3M)

### O povo

(111) “A gente discutia, mais jovens, **o povo** discutia mas eu sei, não gostava, mas fomos sempre amigos.” (P08.M3B)

### O sujeito

(112) “É, é perigoso e... a pessoa tem que ter mais cuidado, tem que ter muito cuidado mesmo. E se está com um parceiro tem que ser mesmo ele só, porque se está a desviar, [ININT] mas tá com aquela pessoa, **o sujeito** apanhar mesmo.” (P10.F3B)

### A criança

(113) “porque **a criança** ao ir pra escola os pais ou encarregado de educação devem ter uma conversa com a criança, pra ela saber o porque, quais são os benefícios que ela ganha indo pra escola e essa conversa ajuda muito.” (P24.F1S)

### Homem

(114) “Mas admiro muito as pessoas que têm e quando chegar a altura, se quiser, vou ter. Como sou jovem, como sou homem, **homem** não tem que se preocupar com isso, as mulheres, se calhar, é que têm. (P22.M1S)

### Fulana

(115) “deixar a encomenda que elas podem fazer se aquela **fulana** vai sair aquela hora.” (P06.F2B)

Para entender um pouco mais os contextos em que esses sintagmas nominais apareceram, os dados foram submetidos a uma nova rodada assumindo como regra de aplicação a forma nominal. O *input* foi de 0.005, o nível de significância foi 0.037 e o *log likelihood* foi -109.846. Nessa perspectiva, o programa selecionou quatro condicionamentos relevantes, nesta

ordem: grau de indeterminação, inclusão do falante, manutenção/mudança no referente e tempo e modo verbal.

O grupo denominado “grau de indeterminação” no uso da estratégia da indeterminação do sujeito pelos falantes não atua da mesma forma, apresentando, assim, uma diferença nos usos da indeterminação seja pela possibilidade de recuperação do referente dentro do contexto, ... Destaca-se ainda que a forma “a gente” mesmo tendo uma configuração igual às FNs, foi considerada como um pronome, visto que a forma já está cristalizada.

**Tabela 23** - Influência do grau de indeterminação no uso da estratégia *formas nominais* em relação às demais estratégias

<b>Grau de indeterminação</b>	Ocorrências	%	Peso Relativo
Indeterminação parcial/referência explícita	3/8	37.5%	<b>0.881</b>
Indeterminação parcial/referência implícita	17/745	2.3%	0.393
Indeterminação completa	39/158	24.7%	<b>0.876</b>
TOTAL	59/911	6.5%	<i>Input: 0.005</i>
Nível de significância = 0.037			

Fonte: Elaboração própria.

Os tipos de indeterminação que mais ocorreram foram a indeterminação parcial com referência explícita no contexto com P.R. 0.88 e, logo em seguida, a indeterminação completa com P.R. 0.87. O tipo que compõe o primeiro grupo envolve um grupo de pessoas de maneira específica que exclui os interlocutores da referida ação, como no exemplo:

(116) “É difícil, mas com a dedicação **a pessoa** consegue”. (P29.F2S)

O segundo tipo já expande para outras pessoas do discurso de uma forma mais genérica. Nesse contexto, o falante ao referenciar, estende para um grupo muito maior de pessoas, conforme se vê, a seguir:

(117) “Mas admiro muito as pessoas que têm e quando chegar a altura, se quiser, vou ter. Como sou jovem, como sou homem, **homem** não tem que se preocupar com isso, as mulheres, se calhar, é que têm. (P22.M1S)

Cada vez mais, estudos como os de Menon (1994), Setti (1997), Godoy (1999) e Carvalho (2010) e Assunção (2012) têm apontado a presença dos sintagmas nominais como forma de grande significância na variedade brasileira. Vale destacar que como as formas nominais de alguma forma possui um elemento explícito, o programa gerou o *singleton group*, deste modo, a variável foi descartada pois seus sujeitos foram realizados.

No *corpus* de Luanda, a segunda variável selecionada pelo programa foi a influência da inclusão do falante. Além da indeterminação completa, o peso relativo foi maior quando o falante não se incluía no contexto discursivo com 0.85, como é o caso do termo “o pessoal”, enquanto o a inclusão do falante não foi uma variante favorecedora tendo um P.R. de 0.39, como se pode ver nos exemplos, logo após a tabela 24.

**Tabela 24** - Influência da inclusão do falante no uso da estratégia *formas nominais* em relação às demais estratégias

Inclusão do falante	Ocorrências	%	Peso Relativo
Sim	32/738	4.3%	0.398
Não	27/173	15.6%	<b>0.853</b>
TOTAL	59/911	6.5	<i>Input: 0.005</i>
Nível de significância = 0.037			

Fonte: Elaboração própria.

(118) “[...] levam os passageiros de uma forma imprópria, é... daí **o pessoal** manda parar o táxi”. (P21.F3S)

A manutenção/mudança do referente foi a terceira variável que também se mostrou relevante para as FNs e identificou-se que apenas quando se tinha o referente igual da estrutura antecedente não foi favorecedora. As formas isoladas, a 1ª forma de uma série e o referente diferente da estrutura anterior, foram mais decisivas para este grupo com PRs de 0.794, 0.689 e 0.684, nesta ordem.

**Tabela 25** - Influência da mudança/manutenção do referente no uso da estratégia *formas nominais* em relação às demais estratégias

Mudança/Manutenção do referente	Ocorrências	%	Peso Relativo
Referente igual à estrutura anterior	3/279	1.1%	0.121

Referente diferente da estrutura anterior	24/332	7.2%	<b>0.684</b>
1ª forma de uma série	18/193	9.3%	<b>0.689</b>
Forma isolada	14/107	13.1%	<b>0.794</b>
TOTAL	59/911	6.5%	<i>Input: 0.005</i>
Nível de significância = 0.037			

Fonte: Elaboração própria.

(119) “É difícil mas com a dedicação **a pessoa** consegue”. (P29.F2S)

(120) “**Nós** vivemos quase um ano no *Cuito* e era insuportável. **A pessoa** ter pra contar né, corpo arrepiá, nem penso... parece que **a pessoa** que passou mesmo o próprio tempo.” (P29.F2S)

(121) “[...] **você** tá a falar pra ela - uma de 16 anos - fala pra ela e ela só te olha "o que que a mãe tá a dizer?" Até **a pessoa** admira "mas você tá a crescer aonde?" (P29.F2S)

**Tabela 26** - Influência do tempo e modo verbal no uso da estratégia *formas nominais* em relação às demais estratégias

<b>Tempo e modo verbal</b>	Ocorrências	%	Peso Relativo
Presente do Indicativo	48/608	7.9%	<b>0.569</b>
Pretérito Imperfeito do Indicativo	4/135	3.0%	<b>0.518</b>
Pretérito Perfeito	1/107	0.9%	0.157
TOTAL	53/850	6.2%	<i>Input: 0.005</i>
Nível de significância = 0.037			

Fonte: Elaboração própria.

(122) “Se pudesse, há alguns... a mim autocarros a operar na cidade, seria bom se tivesse mais autocarro desse gênero, porque eles... porque **a pessoa** *pode sair* do 1º de Maio até a Viana paga apenas 5 *cuanzas*.” (P31.F3S)

(123) “parece que **a pessoa** que *passou* mesmo o próprio tempo, que *sofreu* mesmo [...]” (P29.F2S)

(124) “mas eu acho que não é por causa da bebida, porque ela não bebia muito, a minha irmã (inint) **a pessoa** dizia[...].” (P07.F2M)

## 6.4 FORMAS PRONOMINAIS X NÃO PRONOMINAIS

Com base nas estratégias mais frequentes nesta pesquisa, ainda foram feitas duas rodadas que confrontassem as formas “pronominais” (“a gente”, “nós” e “você”) *versus* as “não pronominais” ( $\emptyset$ +V3PS”, “ $\emptyset$ +V3PP”, “ $\emptyset$ +V+Se”, “ $\emptyset$ VINF” e as formas nominais). A primeira dessas duas rodadas incluía apenas os dados da primeira pessoa do plural (nós). Já, na segunda, a quantificação foi realizada com todos os dados que faziam referência à primeira pessoa do plural (nós +  $\emptyset$ V1PP”). Buscou-se com essas rodadas observar se os falantes de Luanda utilizavam para indeterminar o sujeito mais as formas pronominais plenas ou as flexões verbais.

No quadro 16 a seguir apresentam-se os pesos relativos das formas citadas.

**Quadro 16** - Pronominais X não-pronominais com dados apenas de “nós”

Inclusão do falante	Sim: 0.653 Não: 0.069
Forma antecedente/ Paralelismo	Forma precedida de <i>a gente</i> explícito: 0.327 Forma precedida de <i>eles</i> explícito: 0.826 Forma precedida de <i>nós</i> explícito: 0.711 Forma precedida de <i>nós</i> implícito: 0.918 Forma precedida de <i>tu</i> explícito: 0.206 Forma precedida de <i>você</i> explícito: 0.528 Forma precedida de <i>você</i> implícito: 0.316 Forma precedida de FN explícito: 0.104 Forma precedida de $\emptyset$ +V+SE: 0.075 Forma precedida de $\emptyset$ +V1PP: 0.710 Forma precedida de $\emptyset$ +V2PS: 0.079 Forma precedida de $\emptyset$ +V3PS: 0.248 Forma precedida de $\emptyset$ +V3PP: 0.307 Forma precedida de $\emptyset$ +VINF: 0.374
Preenchimento do sujeito	Sujeito realizado: 0.799 Sujeito não realizado: 0.294
Tipo de verbo	<i>Dicendi</i> : 0.201

	Ação: 0.441 Epistêmicos: 0.554 Estado: 0.738
Grau de indeterminação	Indeterminação parcial/referência explícita: 0.464 Indeterminação parcial/referência implícita: 0.575 Indeterminação completa: 0.194
Mudança/Manutenção do referente	Referente igual à estrutura anterior: 0.658 Referente diferente da estrutura anterior: 0.357 1ª de uma série: 0.525 Forma isolada: 0.515
Tempo e modo verbal	Presente do Indicativo: 0.497 Pretérito Perfeito: 0.564 Pretérito Imperfeito do Indicativo: 0.368 Futuro do Presente: 0.862 Futuro do Subjuntivo: 0.838
Escolaridade	Ensino Fundamental: 0.537 Ensino Médio: 0.565 Ensino Superior: 0.422

Fonte: Elaboração própria.

*Input*:0.015

Nível de significância: 0,010

As variáveis selecionadas pelo programa *GoldvarbX* são apresentadas pela ordem de relevância estabelecida no Quadro 17.

**Quadro 17 - Pronominais X não-pronominais com dados de “nós” + “Ø+V1PP”**

Preenchimento do sujeito	Sujeito realizado: 0.972 Sujeito não realizado: 0.096
Inclusão do falante	Sim: 0.588 Não: 0.188
Grau de indeterminação	Indeterminação parcial/referência explícita: 0.345 Indeterminação parcial/referência implícita: 0.574

	Indeterminação completa: 0.203
Tempo e modo verbal	Presente do Indicativo: 0.474 Pretérito Perfeito: 0.715 Pretérito Imperfeito do Indicativo: 0.375 Futuro do Presente: 0.514 Futuro do Subjuntivo: 0.887

Fonte: Elaboração própria.

*Input*: 0.115

Nível de significância: 0.010

Na tabela 27, a seguir, constam as variáveis selecionadas pelo programa partindo das rodadas realizadas e de sua ordem de seleção para o “nós”, estratégia indeterminadora mais frequente utilizada em Luanda por meio do confronto com as demais variantes. A tabela permite verificar mais facilmente o comportamento dos grupos mais favorecedores para o fenômeno analisado bem como aqueles que não foram relevantes e que mais inibiram:

**Tabela 27** – Comportamento da estratégia “nós” diante das demais variantes

	A gente	Você	Ø+V3PS	Ø+V3PP	Ø+V+SE	Ø+VINFIN	Formas nominais
<i>Grau de indeterminação</i>		6º			7º		2º
<i>Função da Indeterminação</i>	2º		5º				
<i>Preenchimento do sujeito</i>		2º	2º	2º	3º		3º
<i>Tipo de frase</i>							
<i>Tipo de Verbo</i>		4º		4º	6º		6º
<i>Transitividade Verbal</i>							
<i>Núcleo do predicado</i>							
<i>Tipo de Oração</i>					4º		7º
<i>Tempo e modo verbal</i>		3º		3º			5º
<i>Formas nominais do verbo</i>		5º	6º			3º	
<i>Mudança/Manutenção do referente:</i>	4º		4º			4º	
<i>Forma antecedente/paralelismo</i>	3º	1º	3º		2º	2º	
<i>Inclusão do Falante</i>			1º	1º	1º	1º	1º
<i>Ausência/presença preposição (SE)</i>							
<i>Flexão do verbo (SE)</i>							
<i>Posição do argumento interno do verbo (SE)</i>							
<i>Concordância com o argumento interno do verbo (SE)</i>							

<i>Sexo</i>							4º
<i>Faixa Etária</i>							
<i>Escolaridade</i>	1º				5º	6º	
<i>Língua de Interação</i>						5º	

Input:

Fonte: Elaboração própria.

Conforme se vê na tabela 27, percebe-se que não houve seleção para os seguintes grupos pela “nós”: transitividade verbal, núcleo do predicado, ausência/presença preposição (SE), flexão do verbo (SE), posição do argumento interno do verbo (SE) e a concordância com o argumento interno do verbo (SE). Assim, todas as variantes linguísticas que estavam relacionados ao clítico *se*, semelhantemente à tabela que apontava o comportamento da estratégia “Ø+VIPP” diante das demais variantes, não foram relevantes.

Dentre os grupos selecionados, alguns deles foram ficaram mais em evidência do que outros. Os grupos que apresentaram maior recorrência foram os do grupo linguístico: o “tempo e modo verbal”, sendo selecionado pelo “a gente”, “Ø+V3PS”, “Ø+V+SE”, “Ø+VINFINF” e “formas nominais”; as “formas nominais do verbo” selecionadas foram “Ø+V3PS”, “ØVINFINF” e “Ø+V+SE”; a “mudança e manutenção do referente” pelo “você” e pelo “Ø+V3PP”; a forma antecedente/paralelismo teve mais relevância para as variantes e “você” e “Ø+V3PP”; o “tipo de frase”, selecionado pelo “Ø+V3PP”, “Ø+V3PS” e “Ø+V+SE”, nessa ordem; o tipo de oração, teve seu destaque sendo selecionado pelo “ØV3PP”, “Ø+VINFINF” e “a gente”;

Houve, ainda, seleção para o tipo de frase pelo “Ø+V3PS”, “ØV3PP” e “Ø+V+SE”; o núcleo do predicado foi eleito apenas pelo “Ø+VINFINF”. Para as duas variáveis semânticas (grau de indeterminação e função da indeterminação). O “nós” selecionou apenas a função da indeterminação (uma variante de caráter semântico), enquanto a função da indeterminação foi mais relevante para o “nós”, o “você” e “formas nominais”, todas ela na primeira ordem de seleção.

Em relação às variáveis sociais, apenas quatro variantes foram destacadas: o “a gente”, “Ø+V+SE” e “Ø+VINFINF”. A “faixa etária” foi apenas selecionada pelos fatores “a gente”, o “Ø+V+SE” e “ØVINFINF”, respectivamente. A “escolaridade” e o “sexo” foram selecionadas apenas pelo “ØVINFINF”, enquanto, a língua de interação foi selecionada pelas variáveis “formas nominais” e “Ø+V+SE”. Vale destacar que as variáveis sociais apareceram nas últimas colocações diante da análise estatística, com exceção da faixa etária para a forma “a gente”.

Com base nas nestas duas tabelas, as tabelas 04 e 27, referente aos comportamentos em relação às outras variáveis, pela ordem de seleção, nota-se que há muitas similaridades entre as duas, desde a seleção das formas relevantes ou até mesmo das descartadas. Em muitos casos, houve apenas alteração de uma variante ou da sua ordem de posição. Diante disso, conclui-se que as rodadas com dados apenas “ØV1PP” e a amalgamada com “nós” tiveram comportamentos bem parecidos, com pouquíssimas alterações.

## 7 CONSIDERAÇÕES FINAIS

Nesta pesquisa, propôs-se analisar as estratégias de indeterminação do sujeito na fala dos moradores de Luanda (Angola), buscando responder a algumas questões relacionadas aos recursos de indeterminação do sujeito mais utilizados pelos falantes luandenses no uso vernacular e culto, observando se há similaridades com os resultados encontrados na variedade popular oral do PB. Deste modo, foi crucial analisar quais contextos favoreciam o uso das formas não legitimadas pelas gramáticas tradicionais, se os linguísticos ou os sócio-históricos.

O trabalho está vinculado ao projeto de pesquisa “Em busca das raízes do Português Brasileiro: estudos morfossintáticos”, em sua fase III, do Núcleo de Estudos em Língua Portuguesa (NELP) sediado na Universidade Estadual de Feira de Santana (UEFS). Assim, foram utilizadas 32 entrevistas com falantes de Português como L1, estratificadas socialmente da seguinte forma: três faixas etárias em três níveis de escolaridade, sexo e língua de intercomunicação. Para a realização desse estudo, foram elencadas, à princípio, quinze variáveis dependentes, mas, algumas delas foram retiradas pelo número pequeno de ocorrências na amostra.

Foram analisadas sete variantes, ao total, cinco delas tinha sujeito lexical preenchido (*nós, a gente, você, ele, você*); quatro delas, sem grupo lexical preenchido com algumas delas consideradas estratégias canônicas pelas gramáticas tradicionais (o *verbo na terceira pessoa do singular acompanhado da partícula “se” / (Ø+V+SE)*, o *verbo na terceira pessoa do plural / (Ø+V3PP)*, o *verbo no infinitivo (Ø+VINF)* e por fim, o *verbo na terceira pessoa do singular (Ø+V3PS)*; as presenças das formas nominais (FN).

As variáveis linguísticas utilizadas foram grau de indeterminação, função da indeterminação (estas duas de caráter mais semântico), tipo de frase, tipo de oração, flexão do verbo (SE), tempo e modo verbal, transitividade verbal, tipo de verbo, estrutura do núcleo do predicado, ausência e presença de preposição, concordância com o argumento interno do verbo (SE), posição do argumento interno do verbo, preenchimento do sujeito, mudança/manutenção do referente, forma antecedente/paralelismo.

Foi usado o método quantitativo com o suporte do programa estatístico *Goldvarb X* e as análises foram feitas à luz da Teoria da Variação e Mudança LABOV (1972). Para a análise foram feitas algumas rodadas, e algumas variantes precisaram ser desconsideradas para as rodadas, porém comentadas. Ao longo da tese, as análises realizadas se basearam nas nove estratégias mais utilizada pelos luandenses entrevistados: “a gente”, “nós”, “você”, “Ø+V+SE”,

“Ø+V1PP”, “Ø+V3PS”, “Ø+V3PP”, “Ø+VINF” e as formas nominais (sintagmas). A primeira rodada foi feita apenas com os dados de Ø+V1PP (forma flexionada do plural sem sujeito lexicalmente preenchido), no entanto, foram brevemente comentadas, pois, a análise principal uniu essa forma com a estratégia “nós”, visto que em ambos, há referência à mesma pessoa, além de terem sido as mais produtivas da amostra.

Os primeiros resultados do Ø+V1PP apontaram para cada uma rodada com as demais. Na rodada com o Ø+V1PP *versus* Ø+V3PS, os grupos selecionados foram as formas nominais, mudança/manutenção do referente, tempo e modo verbal e tipo de frase; sobre o Ø+V1PP *versus* Ø+V3PP, os grupos foram selecionados na seguinte ordem: forma antecedente/paralelismo; tipo de frase e oração; na análise de Ø+V1PP *versus* Formas nominais, o programa selecionou, diferentemente, duas primeiras, também uma variável social, “língua de intercomunicação”, em terceira posição e, as demais foram a função da indeterminação, a mudança/manutenção do referente e também o tempo e modo verbais. Em “Ø+V1PP” *versus* “A gente”, as variáveis continuam a ser, predominantemente, linguísticas, selecionando além do grau de indeterminação, tempo e modo verbal, tipo de oração, a variável social, faixa etária. As rodadas de “Ø+V1PP” *versus* “você”, e a de “Ø+V1PP” *versus* “nós”, apresentaram, somente, variáveis internas como a função da indeterminação (ambos em primeiro lugar) e as demais mudança/manutenção do referente e forma antecedente/paralelismo, nessa mesma ordem. Ressalta-se que em “Ø+V1PP” *versus* “nós”, só, a “função da indeterminação” foi elencada.

Nas últimas rodadas do “Ø+V1PP” *versus* as demais, nos dados encontrados dessa estratégia com a forma flexionada seguida do clítico, “Ø+V+SE”, já houve um maior número de grupos selecionados. Dos seis grupos, houve o favorecimento dos fatores linguísticos e apenas dois sociais, a apresentadas na seguinte ordem de seleção: formas nominais do verbo; tempo e modo verbal, faixa etária, transitividade verbal, tipo de frase e língua de intercomunicação. Já, na última rodada para o “Ø+V1PP”, quando confrontados com o “Ø+VINF”, apontaram, mesmo nas três últimas posições, um menor favorecimento de três variáveis sociais (escolaridade, faixa etária e sexo), enquanto, as outras quatro foram o núcleo do predicado, as formas nominais do verbo, o tipo de oração e o tempo e modo verbal. A variável escolaridade até então não tinha selecionada em nenhuma das rodadas anteriores, o que pode apontar que, para a forma Ø+V1PP (flexão do plural sem o sujeito lexicalmente preenchido, em relação ao “Ø+VINF”, não exerceu muita influência.

Observa-se, então, que o paralelismo sintático-discursivo, tendo estreita relação com mudança/manutenção do referente, mostrou-se ser uma das variáveis linguísticas mais favorecedoras no uso de diferentes estratégias de indeterminação do sujeito no PA, assim como frequentemente acontece com o PB, sendo selecionado por quatro dos grupos apresentados.

Antes de partir para as rodadas com “ $\emptyset$ +V1PP” e “nós” amalgamados, foi feita uma análise para ver a interação entre essas duas variáveis, isoladamente. Os resultados mostraram valores praticamente invertidos em relação a uma variante à outra. Os pesos relativos foram altos, de 0.931 para o “sujeito não realizado”, quando a regra de aplicação foi o “ $\emptyset$ +V1PP” e 0.979 quando a regra de aplicação foi o “nós”, em que apenas o grupo do preenchimento do sujeito foi selecionado pelo programa, o que já era esperado.

Após amalgamento das estratégias o “ $\emptyset$ +V1PP” *versus* “nós”, o procedimento realizado foi o mesmo do que já foi apresentado. O destaque foi para as “formas nominais”, “ $\emptyset$ +V+SE”, “ $\emptyset$ +VINFINF” e “você” mais grupos de fatores selecionados, os dois primeiros, com sete grupos de fatores, e, os dois últimos, com seis. As duas variantes com menos quantidade de fatores selecionadas foram o “ $\emptyset$ +V3PP” e o “a gente” com quatro fatores relevantes, cada. Não houve nenhuma seleção para o uso da terceira pessoa do singular “ $\emptyset$ +V3PS”, em sua forma apenas flexionada, e, conseqüentemente, sem sujeito preenchido lexicalmente.

Nas rodadas binárias que seguem, a regra de aplicação utilizada foi a “nós”, com 451 total ocorrências da amostra analisada. A primeira, “nós” *versus* “a gente”, o ensino superior foi a única escolaridade que favoreceu, apresentando peso relativo maior de 0.809, seguidos pelos ensinos médio e fundamental. Sobre a influência da indeterminação, a variante “desfocalização do sujeito” foi também a única favorecedora e, o paralelismo ocorreu diante de mais dados precedidos de “ $\emptyset$ +V1PP”, seguido pela “forma precedida de *nós* explícito” próximo ao ponto neutro (0.489). Sobre a “mudança/manutenção do referente” os pesos mais altos foram encontrados quando a estrutura foi diferente da anterior (0.762), seguido do referente igual à estrutura anterior (0.516).

Na rodada binária, “nós” *versus* “você”, a forma antecedente foi a primeira selecionada em que as formas precedidas de “ $\emptyset$ +V1PP” e de “*nós* explícito”, para o grupo de forma antecedente, apresentando 0.801 e 0.565 de peso relativo, respectivamente. O sujeito não realizado foi favorecedor com 0.851 e, a influência do tempo e modo verbal aconteceu mais diante dos pretéritos, imperfeitos e perfeitos, ambos do indicativo na seguinte ordem, 0.875 e 0.796 de P.R. Falando ainda sobre o tipo de verbo no uso indeterminador, os verbos de estado e epistêmicos favoreceram com pesos de 0.868 e 0.631. A forma infinitiva apresentou um leve

favorecimento de 0.561 de peso relativo e “o grau de indeterminação”, o último selecionado pelo *GoldvarbX*, os contextos de indeterminação que tiveram maiores favorecimentos foram quando foi possível identificar, pelo menos um dos seus referentes, por meio de inferência, obtendo um P.R. pouco acima do ponto neutro, 0.536, com 82.7% dos dados.

A outra rodada binária analisada foi a de “nós” versus “formas nominais” e *GoldvarbX* selecionou oito grupos de fatores na seguinte ordem: inclusão do falante, com atenção para a situação de estarem inclusos na referência do sujeito; o grau de indeterminação, com percentuais e pesos maiores diante de indeterminação do sujeito parcial e com referência implícita; sobre o preenchimento do sujeito, o destaque foi para as formas não marcadas lexicamente; os homens utilizaram a estratégia “nós” com mais frequência obtendo um P.R. de 0.655 do que as mulheres, em relação ao sexo; o tempo e modo verbal, mais uma vez, foram indicados os tempos pretéritos tipo de verbo (de ação) e por fim, o tipo de oração com favorecimentos das orações “principal” e “absoluta” na indeterminação do “nós”.

“Nós” versus  $\emptyset$ +V3PS, outra rodada binária realizada, bem similar à rodada anterior, teve a realização do sujeito preenchido e quando o falante se encontrava incluído. O terceiro grupo, “paralelismo/forma precedente” relevou que, grande parte dos dados de “nós”, era precedido por ele mesmo ou quando era precedida pela forma vazia ( $\emptyset$ +V1PP). A mudança/manutenção do referente foi selecionada em formas isoladas (0.793), a primeira de uma série (0.748) ou quando o seu referente é igual à estrutura anterior (0.561).

“Nós” versus “ $\emptyset$ +V3PP”, a variante “sim” para a inclusão do sujeito também foi selecionada, alcançando um PR. de 0.698 e o “sujeito realizado” para o “preenchimento do sujeito, com 0.787 de P.R. Os fatores de tempo e modo verbal, diferente das outras rodadas, foram mais influenciados pelo “futuro do subjuntivo”, pelo “presente do indicativo” e pelo “pretérito perfeito”, com os seguintes pesos: 0.604, 0.596 e 0.532. Os verbos de “estado” e “epistêmico” foram mais favorecedores em relação à “ $\emptyset$ +V3PP” no uso do “nós” indeterminador.

A rodada do “Nós” versus “ $\emptyset$ +V+SE”, selecionou a inclusão do falante pelo fator “sim”; o paralelismo, por meio das formas precedidas de formas nominais explícitas, de “VINFINF” e da estratégia “ $\emptyset$ +V1PP”, com pesos relativos categóricos de 1.000, 0.999 para as duas primeiras e 0.694, para “ $\emptyset$ +V1PP”. O sujeito que prevaleceu foi o realizado e o tipo de oração de destaque foi a principal. Sobre a “escolaridade”, o ensino fundamental ficou bem próximo ao peso relativo máximo (0.991). Sobre o “tipo de verbo”, predominaram os de estado e epistêmicos e

o grau de indeterminação, último grupo selecionado, foi favorecido pelo contexto de indeterminação parcial/referência implícita, com 0.545 de peso relativo.

O “nós” *versus* “Ø+VINFIN”, seis variáveis foram selecionadas, as quatro primeiras linguísticas, paralelismo/forma antecedente, formas nominais do verbo e mudança ou manutenção do referente. Na primeira, a inclusão do falante, o fator sim apresentou 0.549 de P.R.; o paralelismo/forma antecedente foi favorecido pelo “nós” diante de uma forma precedida por uma forma nominal, seguido por uma outra forma precedida pelo “nós explícito; as “formas nominais do verbo” foram marcadas pelo gerúndio e, “a mudança ou manutenção do referente” com os dados de contextos isolados e de estruturas iguais; Sobre as variáveis extralinguísticas, a língua de intercomunicação e a escolaridade foram influenciadas por aqueles que não falavam apenas o português e por aqueles de ensino médio, seguido pelos falantes do ensino superior.

Com as “formas nominais” (sintagmas nominais), também foi feita uma rodada separada das demais e percebeu-se que algumas formas foram usadas pelos falantes para indeterminar o sujeito como foi o caso das expressões “(s) pessoa(s)/o pessoal”, “o sujeito”, “a criança”, “homem” e “fulana”. O programa selecionou quatro condicionamentos: grau de indeterminação, inclusão do falante, manutenção/mudança no referente e tempo e modo verbal. Nesse contexto as referências de indeterminação parcial com referência explícita e indeterminação completa foram as maiores favorecedoras com pesos 0.881 e 0.876, respectivamente, esta última, contribuindo para que a “não inclusão” do falante obtivesse valores altos. A penúltima variável, “a manutenção/mudança do referente” só não apresentou favorecimento do “referente igual à estrutura anterior” com pesos relativos bem abaixo dos demais. Sobre a última variável, o “tempo e modo verbal” foi favorecido pelos presentes do indicativo e pelo pretérito imperfeito do indicativo, com pesos um pouco acima do ponto neutro.

Em seguida, é feito um resumo com as formas “pronominais” *versus* “não pronominais” em relação aos dados de “nós” *versus* “ØVIPP”, com prevalência da primeira, apresentando mais grupos selecionados. Em todas as análises aqui feitas, ficou evidente que as variáveis internas foram as que mais proporcionaram o fenômeno analisado, enquanto, as sociais, raramente foram escolhidas pelo programa.

Destaca-se aqui que as hipóteses levantadas para esta pesquisa sobre a indeterminação do sujeito foram comprovadas já que foi possível visualizar que houve o aparecimento de novas formas linguísticas no PA e atuam, de certa maneira, semelhantes em relação ao mesmo fenômeno linguístico na variedade brasileira.

No que concerne ao fenômeno analisado, há variabilidade no sistema não limitando-se àquelas prescritas pelas GTs. Muito disso se deve, provavelmente, à importância do contato entre os europeus e os povos locais que se deu de maneira bastante latente no continente africano, tendo ainda, como pano de fundo, o sistema escravocrata.

Sobre o “*se*” como indeterminador do sujeito, não se mostrou uma variante tão presente na amostra, embora tenha sido encontrado nos três níveis de escolaridade. Destaca-se que os falantes de menos escolarizados usaram, em menor quantidade, a forma de indeterminação com o clítico, o que se explica o fato dessa estratégia ser aprendida, primordialmente, nos espaços escolares. Assim, a forma mais tradicional foi utilizada basicamente por falantes do ensino médio e superior, ao passo que as variantes mais inovadoras estão sendo, cada vez mais, utilizadas por todos os grupos, principalmente, por falantes menos escolarizados, o que pode gerar, ao longo do tempo, um possível desuso da forma canônica “ $\emptyset+V+SE$ ”, influenciando o falar das próximas gerações. Por ora, o fenômeno da indeterminação do sujeito tem se mostrado passar por um processo de variação estável, podendo gerar futuramente ou não uma mudança em progresso na fala da capital de Angola.

Este estudo forneceu dados para a compreensão de como aconteceram os contatos linguísticos ocorridos nas áreas em que a língua portuguesa se expandiu, em especial de Angola e Brasil. É categórico nas pesquisas já realizadas sobre a indeterminação do sujeito no PB, o fato de os falantes utilizarem muito mais estratégias de indeterminação de sujeito não consideradas “legítimas” pelas gramáticas normativas, em situações reais de fala vernacular, como atestam Duarte e Reis (2018), em dados de fala de PB culto do Rio de Janeiro, em que houve um resultado preferencial e ascendente da forma preenchida de sujeito, pronominal, “nós”. Nesta pesquisa, ao contrário do PB, os resultados do português falado em Luanda, a variante que apresentou um resultado com maior número de ocorrência foi a forma não preenchida “ $\emptyset V I P P$ ” indeterminador.

Pode-se concluir que é urgente e necessário estimular mais estudos de diversos fenômenos variáveis (não só da indeterminação) que foquem em outras variedades do português, que não sejam limitadas à brasileira ou à europeia. Assim, é de suma importância estreitar os laços por meio de pesquisas que aproximem mais o Brasil e África. Reafirma-se que esta pesquisa não deve se esgotar aqui e que outros estudos possam colaborar em uma análise contrastiva vislumbrando a ampliação de debates com a comunidade científica sobre as realidades sociolinguísticas de Angola e do Brasil.

## REFERÊNCIAS

ALENCASTRO, L. F. **O trato dos viventes: formação do Brasil no Atlântico Sul**. São Paulo: Companhia das Letras, 2000.

ALMEIDA, N. M. de. **Gramática metódica da língua portuguesa**. 44 ed. São Paulo: Saraiva, 1999.

ANGOLA. **Constituição da República de Angola**. Luanda: Assembleia Nacional, 2010. Disponível em <[https://imgs.sapo.pt/jornaldeangola/content/pdf/CONSTITUICAO-APROVADA\\_4.2.2010-RUI-FINALISSIMA.pdf](https://imgs.sapo.pt/jornaldeangola/content/pdf/CONSTITUICAO-APROVADA_4.2.2010-RUI-FINALISSIMA.pdf)>.

ANGOLA. **Lei de Bases do sistema de educação**. Luanda, dez., 2001. Disponível em <[https://www.unicef.org/angola/media/1381/file/Lei%2013\\_01\\_Lei%20de%20Bases%20do%20Sistema%20de%20Educacao%20de%20Angola%202001.pdf](https://www.unicef.org/angola/media/1381/file/Lei%2013_01_Lei%20de%20Bases%20do%20Sistema%20de%20Educacao%20de%20Angola%202001.pdf)>.

ARAUJO, S. S. de F. **Nosso, da gente e de nós: um estudo sociolinguístico da expressão de posse no português rural afro-brasileiro**. Dissertação (Mestrado em Letras) Instituto de Letras, Universidade Federal da Bahia, Salvador, 2005.

ARAUJO, S. S. de F. O uso variável da concordância verbal no português do Brasil (PB) e no português de Angola (PA): A história externa em foco. In: 5º Encontro Açoriano da Lusofonia, 2010, Florianópolis. **Atas/ Anais**, 2010. p. 367-381.

ARAUJO, S. S. de F.; DANTAS, N. dos S. Os verbos ter e haver existenciais no português falado em Luanda-Angola. **Letrônica**, v. 10, p. 64-81, 2017. Disponível em: <https://revistaseletronicas.pucrs.br/letronica/article/view/24809/16320> . Acesso em 04.ago.2025.

ASSUNÇÃO, J. da S. **A indeterminação do sujeito na variedade linguística de Feira de Santana: um estudo variacionista**. Orientadora: Norma Lúcia Fernandes de Almeida. 2012. 119 f. Dissertação (Mestrado em Estudos Linguísticos). Universidade Estadual de Feira de Santana, Feira de Santana-BA, 2012.

ASSUNÇÃO, J. da S.; ALMEIDA, N. L. F. A gente e você: formas de indeterminação do sujeito em Feira de Santana- BA. In: **Variação linguística em Feira de Santana-Bahia**. ALMEIDA, Norma Lúcia Fernandes de; Araujo, Silvana Silva de Farias; TEIXEIRA, Eliana Pitombo; CARNEIRO, Zenaide Oliveira Novais. Feira de Santana: UEFS Editora, 2016, p. 267-309.

AVELAR J; GALVES, C. O papel das línguas africanas na emergência da gramática do português brasileiro. **Revista Linguística** (Online), v. 30, n. 2, p. 241-288, 2014. Disponível em: [https://www.mundoalfal.org/sites/default/files/revista/30\\_2\\_linguistica\\_241\\_288.pdf](https://www.mundoalfal.org/sites/default/files/revista/30_2_linguistica_241_288.pdf). Acesso em: 05.jun.2021.

AZEREDO, J. C. de. **Gramática Houaiss da língua portuguesa**. São Paulo: Publifolha, 2008.

BAGNO, M. **Gramática pedagógica do português brasileiro**. São Paulo: Parábola Editorial, 2011.

BECHARA, E. **Moderna gramática portuguesa**. 10.ed. São Paulo: Companhia Editora Nacional, 1966.

BECHARA, E. **Moderna gramática portuguesa**. 39 ed., rev. e ampl. Rio de Janeiro: Nova Fronteira, 2019.

BENVENISTE, É. **Problemas de linguística geral**. São Paulo: Nacional, 1976.

BENVENISTE, É. **Problemas de linguística geral**. 4. ed., v.1. Campinas, São Paulo: Pontes, Editora da Universidade Estadual de Campinas, 1995.

BERLINCK, R. de A. **A ordem V SN no português do Brasil: sincronia e diacronia**. 1988. 288 f. Dissertação (Mestrado) - Curso de Pós-graduação do Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, Campinas, 1988.

BERNARDO, P. J.; SEVERO, C. G. Políticas linguísticas em Angola: sobre as políticas educativas in (ex)cludentes. **Revista da ABRALIN**, V. XVII, N. 2, 2018.

BERNARDO, P. J.; SEVERO, C. G. Políticas linguísticas em Angola: sobre as políticas educativas in (ex)cludentes. **Revista da ABRALIN**, V. XVII, N. 2, 2018.

BHABHA, H. *O Local da cultura* Belo Horizonte: UFMG, 2003, p. 70.

BLEEK, W.H. I. **A comparative grammar of South African languages**. Part 1. Phonology. London/Trubner: Paternoster R, 1862.

BOAS, F. **Antropologia cultural**. Trad. Celso de Castro. 6. ed. Rio de Janeiro: Jorge Zahar Ed., 2010.

BONVINI, E. O vocabulário de origem africana na constituição do português do Brasil. In: Fiorin, José Luiz; PETTER, M. **África no Brasil: a formação da língua portuguesa**. São Paulo: Contexto, 2008, p.101 -144.

BONVINI, E. Línguas africanas e português falado no Brasil. In: PETTER, M.; FIORIN, J. L. (Orgs.). **África no Brasil: a formação da língua portuguesa**. 2. ed. 4 reimp. São Paulo: Contexto, 2020, p. 15 - 62.

BORTONI- RICARDO, S. M. **Educação em língua materna: a sociolinguística na sala de aula**. São Paulo: Parábola Editorial, 2004.

BORTONI-RICARDO, S. M. **Nós chegemos na escola, e agora?** Sociolinguística e educação. São Paulo: Parábola, 2005.

BORTONI- RICARDO, S. M. **Português Brasileiro: a língua que falamos**. São Paulo: Editora Contexto, 2021.

BOSI, A.. “Plural, mas não caótico” in BOSI, Alfredo (org.). **Cultura brasileira – temas e situações**. São Paulo: Ática, 1987b.

CALVET, L. J. **Tradição oral & tradição escrita**. São Paulo: Parábola Editorial, 2011.

CAREGNATO, L. Domínio colonial português em Angola nos séculos XV e XVI. Associação Nacional de História. In: **X Encontro Estadual de História- ANPUH-RS**. 2010, Santa Maria. Anais do Encontro Estadual de História – ANPUH- RS. Pelotas: Editora da Universidade Estadual de Pelotas, 2010.

CARVALHO, V. de. **Você, a gente et alia indeterminam o sujeito em Salvador**. 2010. 197 f. Dissertação (Mestrado em Estudos da Linguagem) - Universidade do Estado da Bahia, Salvador, 2010.

CASTILHO, A. T. de. **Nova gramática do português brasileiro**. 1. Ed. 5ª reimpressão. São Paulo: Contexto, 2019.

CASTRO, I. **Introdução à história do português: geografia da língua - português antigo**. Lisboa: Edições Colibri, 2004.

CAVALCANTE, S. R. de O. **A indeterminação do sujeito na escrita padrão: a imprensa carioca dos séculos XIX e XX**. 117 f. Dissertação (Mestrado) - Curso de Pós-graduação em Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 1999.

CEGALLA, D. P. **Novíssima gramática da língua portuguesa**. 48. ed. São Paulo: Companhia Editora Nacional, 2008.

CONRAD, R. **Últimos anos da escravatura no Brasil: 1850-1888**. Civilização Brasileira, 1978. p. 45

COUTO, Hildo Honório do. **Linguística, ecologia e ecolinguística: contato de línguas**. 2 ed. São Paulo: Contexto, 2021, p. 31-32.

CROCI, F. A imigração no Brasil. In: MELLO, H.; ALTENHOFEN, C.; RASO, T. (Org.). **Os contatos linguísticos no Brasil**. Belo Horizonte: Editora UFMG, 2011, v. 1, p. 173-186.

CUCHE, D. **A noção de cultura nas ciências sociais**. Bauru-SP: EDUSC, 1999.

CUNHA, C. F. Conservação e inovação no português do Brasil. **O eixo e a roda**, 5. Belo Horizonte, 1986, p. 199-230.

CUNHA, C. de S. (1993). **Indeterminação pronominal do sujeito**. 1993. 79 f. Dissertação (Mestrado) - Pós-Graduação da Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 1993.

CUNHA, C.; CINTRA, L. **Nova Gramática do Português Contemporâneo**. 7. Ed. Lexikon, 2016.

DIAS, V. **A indeterminação do sujeito em textos baianos dos séculos XIX e XX: um estudo sociofuncionalista**. Tese de doutorado – Programa de Pós-Graduação em Língua e Cultura. Salvador: UFBA, 2017.

DUARTE, M. E. L. **A perda do princípio ‘Evite Pronome’ no português brasileiro**. Campinas: UNICAMP, 1995 (Tese de Doutorado).

DUARTE, M. E.; PAIVA, M. da C. Mudança Linguística: observações no tempo real. In: MOLLICA, M. C.; BRAGA, M. L. (Org.). **Introdução à sociolinguística: o tratamento da variação**. 4. ed. São Paulo: Contexto, 2010, p. 179-190.

DUARTE, M.E.L.; REIS, E. P. R. **Revisitando o sujeito pronominal vinte anos depois**. ReVel, São Paulo, vol. 16, 30, p. 173-197, 2018.

ELIA, S. **A unidade linguística do Brasil**. Rio de Janeiro: Padrão, 1979.

FAGE, J. D. **História da África**. Lisboa: Edições 70, 2010.

FANON, F. **Pele Negra, Máscaras Brancas**. Tradução Maria Adriana da Silva Caldas. Rio de Janeiro: Fator, 2008.

FARACO, C. E.; MOURA, F. M. **Gramática**. 10. ed. São Paulo: Editora Ática, 1997.

FONSECA, M. N. S. Literatura e oralidade africanas: mediações. **Revista Mulemba/Revista do Setor de Letras Africanas de Língua Portuguesa** - Departamento de Letras Vernáculas. Faculdade de Letras da Universidade Federal do Rio de Janeiro. v. 14, n. 2, jul-dez de 2016, p.12-34. Disponível em <<https://revistas.ufrj.br/index.php/mulemba/>>.

FOUCAULT, M. **Microfísica do Poder**. São Paulo: Graal, 2016.

FRANCESCHINI, L. T.. **Variação pronominal nós/a gente e tu/você em Concórdia – SC**. Tese (Doutorado em Letras) - Setor de Ciências Humanas, Letras e Artes, Universidade Federal do Paraná. Curitiba, 2011. 252 f.

GALVES, C. O papel das línguas africanas na formação do português brasileiro: (mais) pistas para uma nova agenda de pesquisa. **Gragoatá**, Niterói, v. 24, 2008, p. 143-164.

GIVON, T. **On understanding grammar**. New York: Academic Press, 1979.

GODOY, M. A. M. de. **A indeterminação do sujeito no interior paranaense: uma abordagem sociolinguística**. 1999. 128 f. Dissertação (Mestrado em Letras e Linguística) - Universidade Federal do Paraná, 1999.

GONÇALVES, P. Panorama gérai do Português de Moçambique. In: **Revue belge de philologie et d'histoire**, tome 79, fasc. 3, Langues et littératures modernes - Moderne taal- en letterkunde, 2001, p. 69.

GONÇALVES, P. **A Génese do Português de Moçambique**. Lisboa: INCM - Imprensa Nacional Casa da Moeda, 2010.

GONÇALVES, R; HAGEMEIJER, T. O português num contexto multilingue: o caso de São Tomé e Príncipe. **Revista Científica da Universidade Eduardo Mondlane**: Série Letras e Ciências Sociais 1:1, 2015, p. 87-107.

GREENBERG, J.H. **Studies in African Linguistic Classification**: New Haven: Compass Publishing Company, 1955.

GREENBERG, J.H. **The languages of Africa**. Bloomington: Indiana University Center in Anthropology, Folcklore and Linguistics: The Hague: Mouton, 1963.

GUIMARÃES, T. F. V. **Formas de indeterminação do sujeito no estado da Bahia**: um estudo geo-sociolinguístico. 2018. 99 f. Dissertação (Mestrado em Estudo de Linguagens). Departamento de Ciências Humanas – Campus I- Universidade Estadual da Bahia, 2018.

GUTHRIE, M. **Comparative Bantu**: an introduction to the comparative linguistics and prehistory of the Bantu languages. v. 4. Farnborough: Gregg International, 1967-1971.

GUY, G. **Linguistic variation in Brazilian Portuguese**: aspects of the phonology, syntax, and language history. Doctoral Dissertation. University of Pennsylvania. Ann Arbor: University Microfilms International, 1981.

GUY, G. On the nature and origins of popular brazilian portuguese. In: GUY, G. **Estudios sobre el español de América y lingüística Afroamericana**. Bogotá: Instituto Caro y Cuervo, 1989, p. 227-245.

GUY, G. R.; ZILLES, A. M. **Sociolinguística quantitativa**: instrumental de análise. São Paulo: Parábola, 2007.

HAGEMEIJER, T. O português em contacto em África. In A. M. Martins & E. Carrilho (eds.). **Manual de Linguística Portuguesa**, 43-67. Berlim: Mouton de Gruyter, 2016.

HAUY, A. B. **Da necessidade de uma gramática-padrão da língua portuguesa**. 2.ed. São Paulo: Ática, 1986.

HOBBSAWM, E. A invenção das tradições. Tradução Celina Cardim Cavalcante. HOUAISS, A. **O português no Brasil**. Rio de Janeiro: EBRADF, 2015.

HOLM, J. Creole influence on popular Brazilian Portuguese. In: GILBERT, G. (Org.). **Pidgin and creole languages**: essays in memory of john E. Reineck. Honolulu: University of Hawaii Press, 1987. p 406-429.

HOLM, J. Popular Brazilian Portugueses: a semi-creole. In: D'ANDRADE, E.; KIHN, A. (Org.). **Actas do colóquio sobre crioulos de base lexical portuguesa**. Lisboa: Edições Colibri, 1992. p. 37-66.

HOLM, J. A semicrioulização do português vernáculo do Brasil: evidência de contacto nas expressões idiomáticas. **Papia**, Revista de Crioulos de Base Ibérica, Brasília: Universidade de Brasília/Thesaurus, v. 3, n. 2, 1994.

HOLM, J. **Languages in contact**: the partial restructuring of vernaculars. New York: Cambridge University Press, 2004.

INE, Instituto Nacional de Estatística. **Resultados definitivos do recenseamento geral população e da habitação de Angola 2014**. Luanda, 2016.

INVERNO, L. A transição de Angola para o português: uma história sociolinguística. In: TORGAL, L; PIMENTA, F; SOUSA, J. **Comunidades imaginadas: nação e nacionalismos em África**. Coimbra, 2008, p. 117-129.

INVERNO, L. Contato linguístico em Angola: restrospectiva e perspetivas para uma política linguística. In: PINTO, P; MELO-PFEIFER, S. **Políticas Linguísticas em Português**. Lisboa: Lidel, 2018, p. 82-105.

KLEIMAN, A. B. Modelos de letramento e as práticas de alfabetização na escola. In: KLEIMAN, A. B. (Org.). **Os significados do letramento: uma nova perspectiva sobre a prática social da escrita**. Campinas, SP: Mercado de Letras, 1995. p. 15-61.

LABOV, W. **Padrões sociolingüísticos**. Tradução Marcos Bagno, Maria Marta Pereira Scherre, Caroline Rodrigues Cardoso. São Paulo: Parábola Editorial, 2008 [1972].

LEITE, A. M. Empréstimos da Oralidade na Produção e Crítica Literárias Africanas. União dos Escritores Angolanos. **Ensaio**. 1998. Disponível em <<https://www.ueangola.com/criticas-e-ensaios/item/206-empr%C3%A9stimos-da-oralidade-na-produ%C3%A7%C3%A3o-e-cr%C3%ADtica-liter%C3%A1rias-africanas>>.

LÉVI-STRAUSS, C. **Raça em História**. Lisboa: Editorial Presença, 2008.

LIMA, R. **Gramática normativa da língua portuguesa**. 49. ed. Rio de Janeiro: José Olympio, 2011.

LOPES, C. R. dos S. **Nós e a gente no português falado culto do Brasil**. Dissertação de Mestrado em Língua Portuguesa. Rio de Janeiro: Universidade Federal do Rio de Janeiro, 1993.

LOPES, C. R. dos S. **A inserção de a gente no quadro pronominal do português: percurso histórico**. Rio de Janeiro: UFRJ, Faculdade de Letras, 1999 (Tese de Doutorado).

LUCCHESI, D. O conceito de transmissão linguística irregular e o processo de formação do português do Brasil. In C. Roncarati & J. Abraçado (eds.), **Português brasileiro: contato linguístico, heterogeneidade e história**. Rio de Janeiro: 7Letras, 2003, 272-284.

LUCCHESI, D. História do contato entre línguas no Brasil. In: LUCCHESI, D; BAXTER, A; RIBEIRO, I. (Org.). **O Português Afro-Brasileiro**. Salvador: EDUFBA, 2009, p. 41-73.

LUCCHESI, D. **Língua e Sociedade Partidas: a polarização sociolinguística do Brasil**. São Paulo: Contexto, 2015.

LUCCHESI, D. Por que a criouliização aconteceu no Caribe e não no Brasil? Condicionamentos sócio-históricos. **Gragoatá**, 24(48), 2019, p. 227-255.

LUCCHESI, D.; BAXTER, A. A transmissão linguística irregular. In: LUCCHESI, D.; BAXTER, A.; RIBEIRO, I (Org). **O Português Afro-Brasileiro**. Salvador: EDUFBA, 2009.

- MACEDO, J. R. **História da África**. 1. ed. 5. reimp. São Paulo: Contexto, 2020.
- MACHADO, M. dos S. **Sujeitos pronominais ‘nós’ e ‘a gente’**: variação em dialetos populares do norte fluminense. 1995. 252 f. Dissertação (Mestrado em Linguística) - Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 1995.
- MAHO, J. **A classification of the Bantu languages: an update of Guthrie’s referential system**. NURSE, D., & PHILIPPSON, G. *The Bantu Languages*. Routledge. 2003. <https://doi.org/10.4324/9780203987926>
- MAKONI, S. African languages as European scripts: The shaping of communal memory. In: NUTTALLAND, S.; COETZEE, C. (Orgs.). **Negotiating the past**: The making of memory in South Africa. Cape Town, Oxford Press: 1998.
- MARCUSCHI, L. A. Repetição. In: JUBRAN, C. A. S. e KOCH, I. G. V. *et al.* (Orgs.). **Gramática do português culto falado no Brasil**: Construção do Texto Falado. Campinas: Ed. da Unicamp, 2006. p. 219-254.
- MARCUSCHI, L. A. **Da fala para a escrita**: atividades de retextualização. 10. ed. São Paulo: Cortez, 2010.
- MATTOS e SILVA, R. V. Para a história do português culto e popular brasileiro. In: T. Alkmim (ed.), **Para a história do português brasileiro**. v. III, São Paulo, Humanitas: 443-464.
- MATTOS e SILVA, R. V. **Ensaio para uma sócio-história do português brasileiro**. São Paulo: Parábola, 2004.
- MATTOS E SILVA, R. V. **O português são dois...**: Novas fronteiras, velhos problemas. São Paulo: Parábola, 2004.
- MATTOS E SILVA, R. V. Diversidade e unidade: a aventura linguística do português. **Revista ICALP**. 2008, p. 1-29.
- MATTOS E SILVA, R. V. **Contradições no ensino de português**: a língua que se fala x a língua que se ensina. 8 ed. São Paulo: Contexto, 2016.
- MATTOS E SILVA, R. V. O português brasileiro. **História da língua portuguesa no Brasil**. Disponível em: < <http://cvc.instituto-camoes.pt/hlp/hlpbrasil/>>. Acesso em: 7 jul. 2022.
- MAURER JR. T.H. (1951B). **Dois problemas da Língua Portuguesa**: o infinito pessoal e o pronome se. São Paulo, Ind. Graf. José Magalhães.
- MELLO, H. R. Formação do português brasileiro sob a perspectiva da linguística de contato. In: Mello, H.; Altenhofen, C.; Raso, T. (Org.). **Os contatos linguísticos no Brasil**. Belo Horizonte: Editora UFMG, 2011, v. 1, p. 173-186.
- MELO, G. C. de. **A língua do Brasil**. Rio de Janeiro: Padrão, 1946.

- MENDONÇA, J. de J.; NASCIMENTO, J. dos S.; "Estratégias de indeterminação do sujeito: polidez e relações de gênero", p. 225-238 . In: FREITAG, R. M. K.; SEVERO, C. G. (Org). **Mulheres, Linguagem e Poder** - Estudos de Gênero na Sociolinguística Brasileira. São Paulo: Blucher, 2015.
- MENON, O. P. da S. **Analyse sociolinguistique de l'indétermination du sujet dans le portugais parle au Brésil, à partir des données du NURC/SP**. 397f. Tese (Doutorado em Linguística) - Université de Paris VII, Paris, 1994.
- MENON, O. P. da S. **A indeterminação do sujeito no português do Brasil: NURC-SP e VARSUL**. In: VANDRESEN, Paulino. *Variação, mudança e contato linguístico no Português da Região Sul*. Pelotas: Educat, 2006. Cap. 7, p. 125-167.
- MILANEZ, W. **Recursos de indeterminação do sujeito**. 149 f. Dissertação (Mestrado em Linguística) - Universidade Estadual de Campinas, Campinas, 1982.
- MILROY, J. Ideologias linguísticas e as consequências da padronização. In: LAGARES, X.; BAGNO, M. (Orgs.). **Políticas da norma e conflitos linguísticos**. São Paulo: Parábola Editorial 2011 [2001]. pp. 49-87.
- MINGAS, A. A. **Interferência do Kimbundu no português falado em Lwanda**. Luanda: Chá de Caxinde, 2000.
- MONTEIRO, J. L. **Pronomes pessoais: subsídios para uma gramática do português do Brasil**. Fortaleza: Edições UFC, 1994.
- MONTEIRO, J. L. **Para compreender Labov**. 2 ed. Petrópolis, RJ: Vozes, 2000.
- NARO, A. J. Crioulização e mudança natural. In: NARO, A. J.. **Estudos diacrônicos**. Petrópolis: Vozes, 1973, p. 97-110.
- NARO, A. J. Modelos quantitativos e tratamento estatístico. In: MOLLICA, M. C.; BRAGA, M. L. (Org.). **Introdução à sociolinguística: o tratamento da variação**. 3 ed. São Paulo: Contexto, 2010. p.15-25.
- NARO, A. J.; SCHERRE, M. M. **Origens do Português Brasileiro**. São Paulo: Parábola Editorial, 2007.
- NDOMBELE, E.; TIMBANE, A. O ensino de língua portuguesa em Angola: reflexões metodológicas em contexto multilíngue. **fólio - Revista De Letras**, 12 (1), 2020.
- NETO, S. da S. **Introdução ao estudo da língua portuguesa no Brasil**. 3 ed. Rio de Janeiro, Presença, 1976.
- NEVES, M. H. de M.. **Gramática de usos do português**. São Paulo: Editora UNESP, 2000.
- NEVES, M. H. de M. **Gramática de usos do português**. 2. ed. São Paulo: Editora UNESP, 2011.
- NGUNGA, A. **Introdução à linguística Bantu**. Maputo: Imprensa Universitária, 2004.

NUNES, J. M. (1990). **O famigerado se**: uma análise sincrônica e diacrônica das construções com se apassivador e indeterminador. 1990. 189 f. Tese (Doutorado) - Curso de Pós-graduação da Linguagem, Universidade Estadual de Campinas, Campinas, 1990.

OCHS, E. Planned and unplanned discourse. In: T. Givon. **Syntax and semantics**, v.12. New York: Academic Press, p. 51-80, 1979.

OLIVEIRA, H. T. de. Língua portuguesa em Angola: silenciamentos, isolamentos e hierarquias. **Revista da Abralín**, v. 17, n. 2, p. 210-233, 2018.

OMENA, N. P. de. 66 p. **As formas de referência da 1.a pessoa no plural**. Rio de Janeiro, 1987.

OMENA, N. P. de. A Referência à Primeira Pessoa no Plural: Variação ou Mudança? In: PAIVA, M da C.; DUARTE, M.E.L (Orgs). **Mudança linguística em tempo real**. Rio de Janeiro: Contra Capa Livraria, 2003, p. 63-80.

OMENA, N. P.; BRAGA, L. M. A gente está se gramaticalizando? In: MACEDO, A. T.; RONCARATI, C.; MOLLICA, M. C. (Orgs.). **Variação e Discurso**. Rio de Janeiro: Tempo Brasileiro, 1996. p.75–84.

PAIVA, M. da C. A variável gênero/sexo. In: MOLLICA, M. C.; BRAGA, M. L. (Org.). **Introdução à sociolinguística: o tratamento da variação**. 4. ed. São Paulo: Contexto, 2010, p. 33-42.

PENNYCOOK, A. **Linguística aplicada dos anos 90**: em defesa de uma abordagem crítica. In: SIGNORINI, I.; CAVALCANTI, M. C. (Org.). **Linguística aplicada e transdisciplinaridade: questões e perspectivas**. Campinas: Mercado de Letras, 2007.

PEREIRA, D.C. Uma análise funcionalista da indeterminação do sujeito no Português Popular falado em São Paulo. Universidade de São Paulo, Brasil, 2006.

PEREIRA, D. C. Uma análise funcionalista da indeterminação do sujeito no Português Popular falado em São Paulo. **Filol. linguíst. port.**, São Paulo, 15(2), p. 475-518, Jan./Jun. 2013.

PERINI, M. A. **Gramática descritiva do português**. 4.ed. 5.imp. São Paulo: Ática, 2002.

PESSOA DE CASTRO, Y. **Falares africanos na Bahia**: um vocabulário afro-brasileiro. Rio de Janeiro: Topbooks, 2001.

PESSOA DE CASTRO, Y. A matriz africana no português do Brasil. CARDOSO, S. A. M.; MOTA, J. A.; MATTOS E SILVA, R. V.. (Orgs.). **Quinhentos Anos de História Linguística do Brasil**. Salvador: Secretaria da Cultura e Turismo do Estado da Bahia, 2006, p. 83-116.

PESSOA DE CASTRO, Y. **Marcas de africania no português do brasil**: o legado negroafricano nas américas. *Africanias.com*, n. 6. 2014.

- PESSOA DE CASTRO, Y. **A influência das línguas africanas no português brasileiro**. Salvador, Secretaria Municipal de Educação-Prefeitura da Cidade de Salvador, 2015, p. 3-12.
- PETTER, M. M. T. Línguas Africanas no Brasil. CARDOSO, S. A. M.; MOTA, J. A.; MATTOS E SILVA, R. V. (Orgs.). **Quinhentos Anos de História Lingüística do Brasil**. Salvador: Secretaria da Cultura e Turismo do Estado da Bahia, 2006, p. 117-142.
- PETTER, M. M. T. Uma hipótese explicativa do contato entre o português e as línguas africanas. **Papia**, v. 17, p. 9-19, 2007.
- PETTER, M. M. T. Aspectos morfossintáticos comuns ao português angolano, brasileiro e moçambicano. **Papia**. v. 19, p. 201-220, 2009.
- PETTER, M. M. T. A influência das línguas africanas no português brasileiro. In: MELLO, H.; ALTENHOFEN, C.; RASO, T. (Org.). **Os contatos linguísticos no Brasil**. Belo Horizonte: Editora UFMG, 2011.
- PETTER, M. M. T. Ampliando a investigação do *continuum* afro-brasileiro de português. **Papia-Revista Brasileira de Estudos Crioulos e Similares**, v. 25, n. 2, p. 305-317, 2015.
- PISSURNO, K. C. S. **A concordância verbal de terceira pessoa do plural na variedade moçambicana do Português: uma abordagem sociolinguística**. 222f. Dissertação (Mestrado em Língua Portuguesa). Rio de Janeiro: Faculdade de Letras/UFRJ, 2017.
- PONTE, V. **A indeterminação do sujeito no português popular do interior do estado da Bahia**. 2008. 125 f. Dissertação (Mestrado em Letras e Linguística) - Universidade Federal da Bahia, Salvador, 2008.
- RIBEIRO, D. **O povo brasileiro: evolução e sentido do Brasil**. São Paulo: Companhia das Letras, 1995.
- ROLLEMBERG, V. *et al.* **Os pronomes pessoais sujeito e a indeterminação do sujeito na norma culta de Salvador**. Estudos Linguísticos e Literários. Salvador, n. 11, p. 53-74, 1991.
- SANKOFF, D.; TAGLIAMONTE, S.; SMITH, E. **GoldVarb X – a multivariate analysis application**. 2005. Toronto: Department of Linguistics; Ottawa: Department of Mathematics. Disponível em: [http://individual.utoronto.ca/tagliamonte/Goldvarb/GV\\_index.htm#ref](http://individual.utoronto.ca/tagliamonte/Goldvarb/GV_index.htm#ref). Acesso em: 20 abr. 2020.
- SANTANA, N. M. O. **A indeterminação do sujeito no português rural do semi-árido baiano**. 2006. 153 f. Dissertação (Mestrado em Língua e Cultura) - Programa de Pós-Graduação em Letras e Linguística do Instituto de Letras, Universidade Federal da Bahia, Salvador, 2006.
- SANTOS, E. Aspectos da língua portuguesa em Angola. **Papia**, n. 28(1), p. 25-49. jan/jun. 2018.
- SCRIBNER, S; COLE, M. **The psychology of literacy**. Cambridge, Mass.: Harvard University Press, 1981.

SETTI, A. C. R. **A indeterminação do sujeito nas três capitais do sul do Brasil**. 1997. 116 f. Dissertação (Mestrado em Linguística) - Universidade Federal do Paraná, Curitiba, 1998.

SEVERO, C. G. **Línguas atuais faladas em Angola**: entrevista com Daniel Perez Sassuco. Florianópolis: NUER – Núcleo de Estudos de Identidades e Relações Interétnicas, 2015.

SILVA, M. C. A. ; ARAUJO, S. A formação da identidade linguística do português falado em Angola: uma revisão bibliográfica e notas sobre a sócio-história. **REVISTA PHILOLOGUS**, v. 26, p. 61-78, 2020.

SILVA, M. C. A. da; SENE, M. G. de; ARAÚJO, S. S. de F. Notas sobre o Português falado em Luanda: um estudo sobre sócio-história e crenças linguísticas. **Miguilim – Revista Eletrônica do Netlli**, Crato, v. 7, n. 2, p. 337-353, mai-ago, 2018. Disponível em:< <http://periodicos.urca.br/ojs/index.php/MigREN/article/view/1652>>. Acesso em: 22 jan. 2021.

SILVA, T. T. da. **Identidade e diferença: a perspectiva dos estudos culturais** /Tomaz Thadeu da Silva (org.), Stuart Hall, Kathryn Woodward. Petrópolis, RJ: Vozes, 2012.

SILVA NETO, S. **Introdução ao estudo da língua portuguesa no Brasil**. 5. ed. Rio de Janeiro: Presença/INL, 1986 [1950].

SILVA NETO, S. **Introdução ao estudo da língua portuguesa no Brasil**. Rio de Janeiro: Presença, 1963[1950].

SILVA NETO, S. **A língua portuguesa no Brasil**: problemas. Rio de Janeiro, Acadêmica, 1960.

SOARES, M. **Linguagem e Escola**: uma perspectiva social. 11. ed. São Paulo: Ática, 1994.

SOUZA, S. S. **Formas de indeterminação do sujeito**: uma comparação entre as comunidades rurais Matinha e Casinhas (BA). 2014. 148 f. Dissertação (Mestrado em Estudos Linguísticos) - Universidade Estadual de Feira de Santana, Feira de Santana, 2014.

SOUZA, S. S.; OLIVEIRA, J. M. de. A variação no uso das estratégias de indeterminação do sujeito no português popular da Matinha-BA. In ALMEIDA, N. L. F. de; CARNEIRO, Z. de O. N. **Variação linguística no semiárido baiano**. Feira de Santana: UEFS Editora, 2014. p. 71-100.

TAMANINE, A. M. B. **A alternância nós/a gente no interior de Santa Catarina**. (Dissertação de mestrado). UFPR: Curitiba, 2002.

TAMANINE, A. M. B. **Curitiba da gente**: um estudo sobre a variação pronominal nós/a gente e a gramaticalização de a gente na cidade de Curitiba. (Tese de Doutorado) UFPR: Curitiba, 2010.

TANNEN, D. “Oral and Literate Strategies in Spoken and Written Narratives”. **Language** 58, no. 1. p. 1–21, 1982. <https://doi.org/10.2307/413530>.

TARALLO, F. Sobre a alegada origem crioula do português brasileiro: mudanças sintáticas aleatórias. In: ROBERTS, I.; KATO, M. (Orgs.). **Português brasileiro: uma viagem diacrônica**. Campinas: Unicamp, 1993, p. 35-68.

TCHIMBOTO, B. **Mais de sete milhões de angolanos falam línguas nacionais**, 2016. Disponível em: <<http://ipol.org.br/mais-de-sete-milhoes-de-angolanos-falamlinguas-nacionais/>>.

TEIXEIRA, E. P. **O pronome você no português de Luanda**. São Paulo: FFLCH-USP, 2008. Disponível em <[https://dlcv.fflch.usp.br/sites/dlcv.fflch.usp.br/files/03\\_8.pdf](https://dlcv.fflch.usp.br/sites/dlcv.fflch.usp.br/files/03_8.pdf)>

TEIXEIRA, E. P.; ALMEIDA, N. L. F. A indeterminação do sujeito no português angolano: uma comparação com o português do Brasil. **Papia**, n. 21(1), 2011, p. 99-111.

TEIXEIRA, E. P. In: TEIXEIRA, E. P.; ARAÚJO, S. S. de F. (Org.) **Diálogos entre Brasil e Angola: o português d'aquém e d'além-mar**. Feira de Santana: UEFS Editora, 2017.

TEYSSIER, P. **História da língua portuguesa**. Trad. Celso Cunha. São Paulo: Martins Fontes, 1997.

TFOUNI, L. V. **Letramento e alfabetização**. São Paulo: Cortez, 1995.

TIMBANE, A. A. Que português se fala em Moçambique? Uma análise sociolinguística da variedade em uso. **Vocábulo: Revista de letras e linguagens midiáticas**. v.7, n.1, s.p. 2014, p. 67.

TORRES, P. F. J. ; ARAUJO, S. S. de F. Reflexões sobre oralidade, escrita e ensino em sociedades multilíngues: Angola em foco. In: ANECLETO, Úrsula Cunha; SILVA, Lucas Nascimento. (Org.). **Letramentos e estudos discursivos: práticas, ensino e mídias**. Feira de Santana: UEFS Editora, 2024, v. 1, p. 127-156.

TRUDGILL, P. **Sociolinguistics: an introduction to language and society**. 4. ed. England: Pinguin Books, 2000.

VEIGA, M. **A sementeira**. Portugal: Editor ALAC, 1994.

VEIGA, I. P. A. Projeto Político- Pedagógico e a gestão democrática. Novos marcos para a educação de qualidade. **Revista Retratos da Escola**, Brasília, Vol. 3, 2009. Disponível em: <http://retratosdaescola.emnuvens.com.br/rde/article/viewFile/109/298>. Acesso em 24 nov. 2021.

VOTRE, S. J. Relevância da variável escolaridade. In: MOLLICA, M. C.; BRAGA, M. L. (Org.). **Introdução à sociolinguística: o tratamento da variação**. 4. ed. São Paulo: Contexto, 2010, p. 51-58.

WEINREICH, U.; LABOV, W.; HERZOG, M. **Fundamentos empíricos para uma teoria da mudança linguística**. Trad. Marcos Bagno. São Paulo: Parábola Editorial, 2006 [1968].

ZAU, D. G.. (2011). **A língua portuguesa em Angola: um contributo para o estudo da sua nacionalização**. Tese de Doutorado, Universidade de Beira Interior, Covilhã, 2011.

<file:///C:/Users/alwap/Downloads/Filipe%20Zau%20-%20Tese%20de%20Doutoramento.pdf>

## APÊNDICE A – TABELA DE CODIFICAÇÃO

## TABELA DE CODIFICAÇÃO

<u>Grupo de fatores</u>	<u>Variantes</u>	<u>Códigos</u>
G1: Variáveis dependentes	A gente Eles Eu Tu Nós Você Ø+V+SE Ø+V1PP Ø+V2PS Ø+V3PP Ø+V3PS Ø+VINF VPSA SFO Formas Nominais	g l e t n v \$ * z # & i s % @
G2: Sexo	Masculino Feminino	M F
G3: Faixa Etária	Faixa etária 1 Faixa etária 2 Faixa etária 3	1 2 3
G4: Escolaridade	Ensino Fundamental Ensino Médio Ensino Superior	f m s
G5: Língua de Intercomunicação	Fala apenas português Fala português e compreende/fala bantu	1 2
G6: Grau de Indeterminação	Indeterminação parcial/referência explícita Indeterminação parcial/referência implícita Indeterminação completa	P q c
G7: Função da Indeterminação	Desfocalização do sujeito Exemplificação Descomprometimento Ocultação do sujeito Economia Linguística	K E D O L
G8: Preenchimento do sujeito	Sujeito realizado Sujeito não realizado	e i

G9: Tipo de Frase	Afirmativa Negativa Interrogativa	A B C
G10: Tipo de Verbo	<i>Dicendi</i> Epistêmicos Estado Ação Não se aplica (demais)	d e o a /
G11: Transitividade Verbal	Ligação Transitivo Intransitivo	L T I
G12: Núcleo do predicado	Simplex Composto	s c
G13: Tipo de Oração	Absoluta Coordenada Subordinada Principal	A C S P
G14: Tempo e modo verbal	Presente do Indicativo Pretérito Perfeito Pretérito Imperfeito do Indicativo Futuro do Presente Futuro do Pretérito Presente do Subjuntivo Pretérito Imperfeito do Subjuntivo Futuro do Subjuntivo Imperativo	h d i f c j e u m
G15: Formas nominais do verbo	Infinitivo Particípio Gerúndio Não se aplica	X Y Z /
G16: Mudança/Manutenção do referente:	Referente igual à estrutura anterior Referente diferente da estrutura anterior 1ª de uma série Forma isolada	I D 1 F
G17: Forma antecedente/paralelismo	Forma precedida de a gente explícito Forma precedida de a gente implícito Forma precedida de eles explícito Forma precedida de eles implícito Forma precedida de eu explícito Forma precedida de eu implícito Forma precedida de FN explícito Forma precedida de FN implícito Forma precedida de nós explícito	g G l L e E @ A n

	Forma precedida de nós implícito Forma precedida de você explícito Forma precedida de você implícito Forma precedida de Ø+V+SE Forma precedida de Ø+V1PP Forma precedida de Ø+V2PS Forma precedida de Ø+V3PP Forma precedida de Ø+V3PS Forma precedida de Ø+VINF Forma precedida de VPSA Forma precedida de VPASSINT Forma precedida de Tu implícito Forma precedida de Tu explícito Forma precedida de sujeito oracional	N v V \$ § Z # & I s a T t \
G18: Inclusão do Falante	Sim Não	S N
G19: Ausência/presença preposição (SE)	Presença SE Ausência SE Não se aplica (sem SE)	p a /
G20: Flexão do verbo (SE)	Tempo finito Infinitivo Gerúndio Não se aplica (sem SE)	4 5 6 /
G21: Posição do argumento interno do verbo (SE)	SN posposto SN anteposto	P A /
G22: Concordância com o argumento interno do verbo (SE)	Concordância Não-concordância	C N /