



**UNIVERSIDADE ESTADUAL DE FEIRA DE SANTANA – UEFS**  
Departamento de Letras e Artes  
**PROGRAMA DE PÓS-GRADUAÇÃO EM ESTUDOS LINGÜÍSTICOS**  
**DOUTORADO EM ESTUDOS LINGÜÍSTICOS**

**DAYANE MOREIRA LEMOS**

**ELIHS, UM PROJETO PARA ALÉM DO LITORAL: DA CONSTITUIÇÃO DE  
CORPUS ORAL DE COMUNIDADES AFRO-BRASILEIRAS À ANÁLISE  
ESTATÍSTICA DA VARIAÇÃO NA CONCORDÂNCIA NOMINAL DE NÚMERO NA  
VARIEDADE DA COMUNIDADE AFRO-BRASILEIRA DE VOLTA DO ANGICO  
(BA)**

**ELiHS**

**Estudos Linguísticos e Históricos do Sertão**

Feira de Santana – BA

2023

**DAYANE MOREIRA LEMOS**

**ELIHS, UM PROJETO PARA ALÉM DO LITORAL: DA CONSTITUIÇÃO DE  
*CORPUS* ORAL DE COMUNIDADES AFRO-BRASILEIRAS À ANÁLISE  
ESTATÍSTICA DA VARIAÇÃO NA CONCORDÂNCIA NOMINAL DE NÚMERO NA  
VARIEDADE DA COMUNIDADE AFRO-BRASILEIRA DE VOLTA DO ANGICO**

Tese apresentada ao Programa de Pós-graduação em Estudos Linguísticos (PPGEL), da Universidade Estadual de Feira de Santana (UEFS), como requisito parcial para obtenção do título de Doutora em Estudos Linguísticos.

Orientadora: Prof.<sup>a</sup> Dr.<sup>a</sup> Mariana Fagundes de Oliveira Lacerda

Feira de Santana – BA

2023

**DAYANE MOREIRA LEMOS**

**ELIHS, UM PROJETO PARA ALÉM DO LITORAL: DA CONSTITUIÇÃO DE  
CORPUS ORAL DE COMUNIDADES AFRO-BRASILEIRAS À ANÁLISE  
ESTATÍSTICA DA VARIAÇÃO NA CONCORDÂNCIA NOMINAL DE NÚMERO NA  
VARIEDADE DA COMUNIDADE AFRO-BRASILEIRA DE VOLTA DO ANGICO**

Tese apresentada ao Programa de Pós-graduação em Estudos Linguísticos (PPGEL), da Universidade Estadual de Feira de Santana (UEFS), como requisito parcial para obtenção do título de Doutora em Estudos Linguísticos.

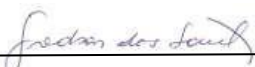
Orientadora: Prof.<sup>a</sup> Dr.<sup>a</sup> Mariana Fagundes de Oliveira Lacerda

Aprovada em 29 de agosto de 2023.



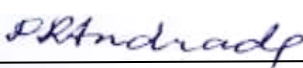
---

Prof.<sup>a</sup> Dr.<sup>a</sup> Mariana Fagundes de Oliveira Lacerda  
Orientadora – Universidade Estadual de Feira de Santana (UEFS)



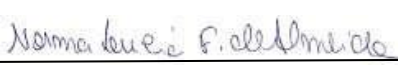
---

Prof.<sup>o</sup> Dr.<sup>o</sup> Gredson dos Santos  
Membro Externo – Universidade Federal da Bahia (UFBA)



---

Prof.<sup>a</sup> Dr.<sup>a</sup> Patrícia Ribeiro de Andrade  
Membro Externo – Universidade do Estado da Bahia (UNEB)



---

Prof.<sup>a</sup> Dr.<sup>a</sup> Norma Lúcia Fernandes de Almeida  
Membro interno – Universidade Estadual de Feira de Santana (UEFS)



---

Prof.<sup>a</sup> Dr.<sup>a</sup> Silvana Silva de Farias Araújo  
Membro interno – Universidade Estadual de Feira de Santana (UEFS)

Ficha Catalográfica - Biblioteca Central Julieta Carteadó - UEFS

L576e

Lemos, Dayane Moreira

ELiHS, um projeto para além do litoral: da constituição de corpus oral de comunidades afro-brasileiras à análise estatística da variação na concordância nominal de número na variedade da comunidade afro-brasileira de Volta do Angico (Ba) / Dayane Moreira Lemos. – 2023.

260 f.: il.

Orientadora: Mariana Fagundes de Oliveira Lacerda.

Tese (doutorado) – Universidade Estadual de Feira de Santana,  
Programa de Pós-Graduação em Estudo Linguísticos, Feira de Santana, 2023.

1. Português popular. 2. Comunidades afro-brasileiras. 3. Sociolinguística.  
4. Interiorização. I. Lacerda, Mariana Fagundes de Oliveira, orient.  
II. Universidade Estadual de Feira de Santana. III. Título.

CDU 81'272 (814.22)

*Dedico às comunidades afro-brasileiras do  
Território de Identidade de Irecê e aos  
monitores do Projeto ELiHS.*

## AGRADECIMENTOS

*“Você não sabe o quanto eu caminhei  
Pra chegar até aqui  
Percorri milhas e milhas antes de dormir  
Eu nem cochilei  
Os mais belos montes escalei  
Nas noites escuras de frio chorei.”  
Cidade Negra<sup>1</sup>*

A finalização da tese desvela uma caminhada, muitas escaladas, que são oprimidas nos muitos sintagmas construídos diariamente na tentativa de chegar a este momento de apresentação do saber científico, da nossa pequena contribuição aos avanços no mapa sociolinguístico do Português Brasileiro. Nestes 04 (quatro) anos aprendi da forma mais encantadora que a ciência se faz presente além dos muros das Universidades, que ser uma doutora em Linguística vai muito além de escrever algumas centenas de páginas, mas se faz na atuação ativa em comunidade, no que deixamos enquanto produto social e científico – e assim que o Projeto ELiHS foi sonhado, implementado e executado no Território de Irecê.

Na transição do sonho à realidade muitas pessoas se fizeram diretamente presentes na vida desta filha, mãe, esposa, professora universitária (60hrs), estudante de doutorado e tantas outras funções que vamos assumindo ao longo da missão de SER MULHER... e por eles chego a este momento dos AGRADECIMENTOS.

Primeiramente a Deus por me conceder a permanência da vida, enquanto milhares de pessoas morriam por conta da COVID, por reger minha vitalidade enquanto cada novo desafio aprecia, por me amparar e conceder forças para que eu pudesse chegar a finalização desta tese.

Palavras me faltam para agradecer aos meus pais, Domingos e Lourdes, pela base familiar, apoio e amor incondicional. Ao meu esposo, companheiro e amigo, Nelson Júnior, que com muito amor e generosidade esteve sempre presente, me compreendendo, apoiando e incentivando nos momentos mais difíceis, quando tudo parecia não mais fazer sentido. Às minhas filha Júlia e Lívia (que assiste a defesa da mamãe na barriga) que me movem, me fortalecem e me fazem sempre seguir em frente. A esses meu eterno amor e gratidão!

Dizem que amigos são famílias que Deus nos permite escolher, por isso gostaria de registrar aqui o nome da Professora Mariana Fagundes Lacerda, que sempre foi muito mais que uma orientadora, foi uma amiga, uma conselheira... tenho certeza que se ela não tivesse sido

---

<sup>1</sup> Compositores: Bino Farias, Lazão, Paulo Gama e Toni Garrido.

tão paciente e compreensiva eu não teria chegado até aqui. Mari, você é um espelho na vida dos seus alunos/orientandos, nos mostra diariamente que ser professor é educar com amor e afeto.

Meus agradecimentos se estendem à Professora Patrícia Andrade que lá em 2008, na graduação em Letras – *Campus V*, me apresentou o mundo da pesquisa científica, me preparando para o mundo da pós-graduação e com o encontro mais efetivo com a pesquisa. Falar em pesquisa é impossível não citar o nome da professora Zenaide Carneiro, que apresentou em 2012 muitas possibilidades de desenvolvimento de pesquisa e o projeto CE-DOHS, que me encantou e me inspirou na implementação do Projeto ELiHS.

Não poderia deixar de agradecer a minha família ELiHS, hoje constituída por 13 integrantes. A colaboração gradual de todos deu vida a um sonho desta doutoranda, me motivou e me fez acreditar que havia um compromisso muito maior do que um dia eu pudesse mensurar, um compromisso com a formação de cada um dos monitores, com líderes das comunidades afro-brasileiras que confiaram a mim informações referentes às suas particularidades.

Agradeço também a FAPESB que por meio de seu financiamento foi possível chegar até algumas comunidades afro-brasileira e organizar o material em plataforma *online* gratuita. À Coordenação do Programa de Pós-graduação em Estudos Linguísticos (PPGEL), pela oportunidade dada de sermos alunos da primeira turma de doutorado. Aos amigos que encontramos nesta jornada itinerante, em especial Elias dos Santos, que colaborou imensamente no desenvolvimento da análise estatística através da plataforma R e com sua paciência me fazer compreender um mundo que iria além do GOLDVARB X. Finalmente, agradeço a todos que contribuíram direta ou indiretamente para a realização desta pesquisa e que não foram aqui mencionados.

*“Não podemos mais viver assim. Temos direito à terra. Somos quilombolas” (Vieira Júnior, 2019)*



## RESUMO

Esta tese volta-se ao estudo do Português Brasileiro, especificamente do português popular, visto aqui sob uma perspectiva sócio-histórica (Mattos e Silva, 2008), compreendendo, de modo geral, as mudanças provocadas pelo contato linguístico. Cientes de que a compreensão da realidade social e linguística do interior baiano, bem como de sua formação histórica, constitui um vasto campo de estudos ainda a ser, acertadamente, investigado por estudiosos, objetivamos, no âmbito desta tese, implementar o Projeto Estudos Linguísticos e Históricos do Sertão (ELiHS) na Universidade do Estado da Bahia (UNEB), *Campus XVI – Irecê*, compor e disponibilizar, em plataforma *online* gratuita, *corpus* oral representativo do português popular de comunidades afro-brasileiras, localizadas no Território de Identidade de Irecê, interior da Bahia, favorecendo a construção de uma fundamentação empírica a variedades do português ainda não catalogadas no campo da Sociolinguística e colaborando para a confirmação da hipótese de que o português afro-brasileiro é uma variedade particularmente afetada pelo contato entre línguas, em sua formação histórica (Luchessi, 2009). Na tentativa da conciliação de aspectos de natureza sócio-histórica com aspectos de natureza linguística, esta tese ainda busca empreender um estudo linguístico do material coletado, com foco em aspectos morfossintáticos do português afro-brasileiro, especificamente a concordância nominal de número, na tentativa de que os resultados endossem as pesquisas científicas, em especial as de cunho Sociolinguístico, que estão “para além do litoral”, favorecendo, na medida do possível, a discussão sobre a formação da realidade sociolinguística brasileira. Considerando esse último objetivo, descreveu-se o português afro-brasileiro na comunidade afro-brasileira de Volta do Angico, tentando entender o processo de Transmissão Linguística Irregular, por meio do controle de variáveis linguísticas e socioculturais de 12 (doze) participantes da pesquisa, comparando os resultados com os dados de Andrade (2003) para a comunidade afro-brasileira de Helvécia. A amostra analisada foi constituída por 1253 dados de sintagmas nominais pluralizáveis, analisados à luz da Teoria da Variação e Mudança Linguística (Weinreich; Labov; Herzog, 2006; LABOV, 2008) e submetidos à modelagem estatística de regressão logística com resposta binária, a linguagem de programação denominada R (R CORE TEAM, 2020), por meio de uma de suas interfaces, o *RStudio* (EQUIPE RSTUDIO, 2020), que reportou quantitativamente os condicionadores da escolha da variante de plural padrão nos elementos flexionáveis do sintagma nominal. Os resultados indicam que, na amostra, há uma tendência significativa ao uso não canônico na indicação do plural nos SNs, um resultado estatístico aproximativo aos dados apresentados por Andrade (2003) para Helvécia (89,55% para Volta do Angico e 91% para Helvécia), nos permitindo afirmar que, embora a tendência do português brasileiro seja a marcação explícita de número nos elementos que compõem o SN, ainda há comunidades linguísticas, como as comunidades afro-brasileiras Volta do Angico e Helvécia, Bahia, que não conservam a atual variante prototípica, de marcas explícitas, reafirmando o posicionamento de Lucchesi (2009) sobre a ideia da perda de morfologia flexional no PB popular ser um indicador da influência do contato linguístico. Os resultados e produtos alcançados na finalização desta tese apresentam potencial de impacto, na medida em que colaborarão para a interiorização da descrição linguística, dando visibilidade às comunidades que estão “para além do litoral”, que vivem e resistem no Território de Identidade de Irecê, fazendo-as presentes no mapa da sociolinguística brasileira.

**Palavras-chave:** interiorização; comunidades afro-brasileiras; constituição de *corpus*; estudo morfossintático da concordância nominal de número.

## ABSTRACT

This thesis is focused on the study of the Brazilian Portuguese, especially on the popular Portuguese, seen here under a sociohistorical perspective (Mattos e Silva, 2008), beholding, overall, the changes caused by linguistic interaction. Knowing that the understanding of the linguistic and social reality of bahian interior, as well as its historical formation, constitutes an extensive field of studies yet to be, rightly, investigated by studios, we aim, in this thesis's scope, to implement the Project Linguistic and Historical Study of the Backlands (LHSB - ELiHS in portuguese: Estudo Linguístico e Histórico do Sertão) at the Bahia State University (UNEB), *Campus XVI – Irecê*, to compose and provide, on free online platforms, oral corpus representative of the popular portuguese in the afro-brazilian communities, located in the Territory of Irecê's Identity, bahian interior, favoring the construction of the empirical fundamentation for varieties of portuguese still not cataloged in the field of Sociolinguistic and collaborating into the confirmation of the hypothesis that the afro-brazilian portuguese is a variable particularly affected by the interaction among languages, in its historical formation (Luchessi, 2009). Under the attempt to conciliate the aspects of sociohistorical nature with the ones of linguistic nature, this thesis still aims to undertake a linguistic study of the collected material, with emphasis on morphosyntactic aspects of the afro-brazilian portuguese, especially the number's nominal agreement, in the expectation that the results endorse scientific researches, in special the ones with Sociolinguistic approach, that exemplify a view "beyond the litoral", prioritizing, as possible as it is, the discussion about the formation of the sociolinguistic brazilian reality. Regarding this last objective, the afro-brazilian portuguese was described in the also afro-brazilian community Volta do Angico, trying to understand the process of Irregular Transmission of Language (Transmissão Linguística Irregular), through the control of linguistics and sociocultural variables of 12 (twelve) research participants, comparing the results with the data from Andrade (2003) for the afro-brazilian community in Helvécia. The analyzed sample was constituted by 1253 datas of nominal syntagmas pluralized, analyzed under the Theory of Language Variation and Change (Weinreich; Labov; Herzog, 2006; Labov, 2008) and subjected to the statistical modeling of logistic regression with binary answer, the coding language so-called R (R CORE TEAM, 2020), by means of one of its interfaces, the *RStudio* (EQUIPE RSTUDIO, 2020), that quantitatively reported the choice's conditionals of the standard plural variable in the flexible elements of nominal syntagma. The results show that, in the sample, there is a significative tendency to the non canonic use of the plural indication of SNs, an approximate statistical result to the data presented by Andrade (2003) for Helvécia (89,55% to Volta do Angico and 91% to Helvécia), allowing us to affirm that, although the tendency of the brazilian portuguese is an explicit demarcation of the numerical aspect in the elements that compose the NS, there are still linguistic communities, such as the afro-brazilian communities Volta do Angico and Helvécia, Bahia, that do not conserve the current prototypical variable, of explicit marks, reaffirming Lucchesi (2009)'s standing about the idea of lost in the flexional morphology for the BP popular being an indicator of the influence from linguistic interaction. The results and products reached at the finalization of this thesis present potential to impact, to the extent that will collaborate into the interiorization of linguistic description, bringing visibility to the communities that are "beyond the litoral", that live and resist at the Territory of Irecê's Identity, making them present on the map of brazilian sociolinguistic.

**Keywords:** interiorization, afro-brazilians communities; *corpus* constitution; morphosyntactic study of the number's nominal agreement.

## LISTA DE FIGURAS

<b>Figura 1</b> – Políticas linguísticas – tentativa de extinção da diversidade linguística do território brasileiro .....	35
<b>Figura 2</b> – Línguas cooficiais em municípios brasileiros .....	36
<b>Figura 3</b> – África em 1890 .....	46
<b>Figura 4</b> – Distribuição de Greenberg: troncos linguísticos africanos .....	47
<b>Figura 5</b> – Línguas africanas: troncos linguísticos .....	48
<b>Figura 6</b> – Representação para o conceito de Transmissão Linguística Irregular .....	71
<b>Figura 7</b> – Fases dos estudos da linguagem .....	78
<b>Figura 8</b> – Territórios de Identidade demarcados no Estado da Bahia .....	102
<b>Figura 9</b> – Composição do Território de Identidade de Irecê – BA .....	103
<b>Figura 10</b> – Caracterização do contexto de rurbanização do TII .....	106
<b>Figura 11</b> – Grupos vulneráveis no TII .....	108
<b>Figura 12</b> – Do conceito de Quilombo à comunidade afro-brasileira .....	113
<b>Figura 13</b> – Mapa das Comunidades certificadas pela FCP enquanto CRQ .....	118
<b>Figura 14</b> – Retratos de membros das Comunidades em atividade de campo .....	128
<b>Figura 15</b> – Casa do poço artesiano – Comunidade de Floresta I .....	129
<b>Figura 16</b> – Chafariz de Abastecimento desativado – Comunidade de Segredo .....	129
<b>Figura 17</b> – Cemitério – Comunidade de Cruzeiro .....	130
<b>Figura 18</b> – Lajedo – Comunidade de Brejinho .....	130
<b>Figura 19</b> – Exterior e interior da primeira escola da Comunidade de Floresta II .....	131
<b>Figura 20</b> – Retratos de moradias das Comunidades .....	131
<b>Figura 21</b> – Retratos das religiões predominantes nas Comunidades do TII .....	136
<b>Figura 22</b> – Escola Ensino Fundamental I desativada – Comunidade de Cruzeiro .....	139
<b>Figura 23</b> – Logomarca do Projeto ELiHS: do real à criação .....	146
<b>Figura 24</b> – Logomarca do Projeto ELiHS: fases de constituição .....	147
<b>Figura 25</b> – Organograma da equipe ELiHS: coordenação e subcoordenações .....	148
<b>Figura 26</b> – Metodologia para organização de <i>corpus</i> .....	153
<b>Figura 27</b> – Estrutura hierárquica de responsabilização por pesquisas envolvendo humanos no Brasil .....	155
<b>Figura 28</b> – Plataforma eletrônica CE-DOHS .....	167
<b>Figura 29</b> – Plataforma eletrônica CE-DOHS: <i>corpus</i> eletrônico .....	168
<b>Figura 30</b> – Estrutura geral do Projeto CE-DOHS .....	169

<b>Figura 31</b> – Plataforma eletrônica CE-DOHS: amostras de fala .....	169
<b>Figura 32</b> – ELiHS e CE-DOHS: uma parceria na constituição de <i>corpora</i> orais .....	171
<b>Figura 33</b> – Projeto ELiHS: Rede de Diretório de Pesquisas no Brasil .....	172
<b>Figura 34</b> – Projeto ELiHS: Rede de parceria Territorial .....	173
<b>Figura 35</b> – Tripé tecnológico e científico da Plataforma do Projeto ELiHS .....	175
<b>Figura 35</b> – Página inicial da plataforma do Projeto .....	176
<b>Figura 36</b> – Eixo construtor do Jogo <i>ConcordAção</i> .....	184
<b>Figura 37</b> – Tela inicial do jogo interativo .....	184
<b>Figura 38</b> – Avatares do Projeto .....	185
<b>Figura 39</b> – <i>Layout</i> do jogo de labirinto do <i>ConcordAção</i> .....	187
<b>Figura 40</b> – Mapa de localização da comunidade Volta do Angico – Canarana – Irecê	208
<b>Figura 41</b> – Uso do Aradado manual .....	209
<b>Figura 42</b> – Uso da Bruaca na colheita .....	209
<b>Figura 43</b> – Batida de feijão manual .....	210
<b>Figura 44</b> – Retratos da Comunidade Volta do Angico .....	211
<b>Figura 45</b> – Monitor do Projeto apresentando o Candeeiro .....	211
<b>Figura 46</b> – Antiga cacimba – Volta do Angico .....	211
<b>Figura 47</b> – Atual caixa de abastecimento de água .....	211
<b>Figura 48</b> – Espaço para manutenção da água .....	212
<b>Figura 49</b> – Terreiro de Mamona .....	212
<b>Figura 50</b> – Sopro manual da Mamona .....	212
<b>Figura 51</b> – Plantação de milho .....	212
<b>Figura 52</b> – Criação de suínos .....	212
<b>Figura 53</b> – Escola desativada – Volta do Angico .....	213
<b>Figura 54</b> – Manifestação popular: Festa dos Caretas abril/2023 .....	214
<b>Figura 55</b> – Igreja protestante – Volta do Angico .....	215
<b>Figura 56</b> – Tela de codificação no formato “caso por variável” (Oushiro, 2022) .....	221
<b>Figura 57</b> – Passos metodológico no uso da Plataforma R .....	222

## LISTA DE GRÁFICOS

<b>Gráfico 1</b> – Demografia histórica do Brasil .....	28
<b>Gráfico 2</b> – População censitária (2010) por situação de moradia no TII .....	105
<b>Gráfico 3</b> – Grau de urbanização – Bahia, TI Irecê e municípios do TI (2010) .....	105
<b>Gráfico 4</b> – Proporção de população em extrema pobreza – Bahia, TI Irecê e municípios do TI (2010) .....	107
<b>Gráfico 5</b> – Quantitativo de Comunidades Certificadas e com Regularização Fundiária	126
<b>Gráfico 6</b> – Quantitativo de Associações nas Comunidades do TII .....	133
<b>Gráfico 7</b> – Quantitativo da proporção por grupos etários nas Comunidades do TII .....	134
<b>Gráfico 8</b> – Quantitativo das Comunidades que se autodeclaram enquanto negro e/ou afrodescendentes no TII .....	135
<b>Gráfico 9</b> – Quantitativo das religiões predominantes nas Comunidades do TII .....	136
<b>Gráfico 10</b> – Acesso telefônico nas Comunidades do TII .....	137
<b>Gráfico 11</b> – Acesso à internet nas Comunidades do TII .....	138
<b>Gráfico 12</b> – Níveis de ensino nas Comunidades do TII .....	139
<b>Gráfico 13</b> – Quantitativo de pessoas/membros que concluíram o Ensino Médio nas Comunidades do TII .....	140
<b>Gráfico 14</b> – Resultado contrastivo entre as comunidades Volta do Angico e Helvécia (Andrade, 2003) .....	226
<b>Gráfico 15</b> – Resultado contrastivo entre as comunidades Volta do Angico, Helvécia (Andrade, 2003) e Salvador (Lopes, 2001) .....	227

## LISTA DE QUADROS

<b>Quadro 1</b>	– Línguas subsaarianas .....	26
<b>Quadro 2</b>	– Proposta para periodização da história linguística do Brasil .....	40
<b>Quadro 3</b>	– Traços tipológicos das línguas africanas .....	49
<b>Quadro 4</b>	– Ciclos do tráfico negreiro e origem dos escravos traficados .....	50
<b>Quadro 5</b>	– Periodização da história linguística africana, por Petter (2017) .....	55
<b>Quadro 6</b>	– Concepções de língua .....	84
<b>Quadro 7</b>	– Questões presentes no formulário eletrônico .....	116
<b>Quadro 8</b>	– Mapeamento das Comunidades certificadas pela FCP enquanto CRQ .....	119
<b>Quadro 9</b>	– Passo a passo da Titulação de Território Quilombola .....	125
<b>Quadro 10</b>	– Frentes de trabalho do Projeto ELiHS .....	145
<b>Quadro 11</b>	– Características sociodemográficas das Comunidades que compõem o <i>corpus</i> do Projeto ELiHS .....	159
<b>Quadro 12</b>	– Ficha do participante: influência humana da escolha .....	160
<b>Quadro 13</b>	– Quadro amostral .....	160
<b>Quadro 14</b>	– Cabeçalho padrão da chave de transcrição .....	165
<b>Quadro 15</b>	– Indicações gerais sobre a formatação do texto transcrito .....	165
<b>Quadro 16</b>	– Critérios gerais para orientar a transcrição .....	165
<b>Quadro 17</b>	– CE-DOHS: amostra oral disponível para <i>download</i> .....	170
<b>Quadro 18</b>	– Descrição funcional do <i>menu</i> do <i>site</i> .....	179
<b>Quadro 19</b>	– Acervo da Biblioteca Digital .....	181
<b>Quadro 20</b>	– Estratégia para desenvolvimento das perguntas do jogo .....	186
<b>Quadro 21</b>	– Panorama dos estudos sobre variação na concordância nominal desenvolvidos nas últimas décadas no Brasil (2000 – 2022) .....	199
<b>Quadro 22</b>	– Descrição da população amostra do PPB representativo da Comunidade de Volta do Angico .....	216
<b>Quadro 23</b>	– Descrição das variáveis resposta e preditoras .....	219
<b>Quadro 24</b>	– Descrição das variáveis preditoras linguísticas, após teste <i>qui-quadrado</i> ..	222

## LISTA DE TABELAS

<b>Tabela 1</b>	– Estimativas de população no Brasil .....	27
<b>Tabela 2</b>	– Estimativas de imigração para o Brasil .....	27
<b>Tabela 3</b>	– População de falantes das diversas línguas africanas que chegaram ao Brasil .....	53
<b>Tabela 4</b>	– População escrava .....	54
<b>Tabela 5</b>	– Taxas de uso da concordância no SN em amostra da comunidade Volta do Angico .....	224
<b>Tabela 6</b>	– Estimativas dos parâmetros do modelo (de regressão logística, modelo linear generalizado de efeitos fixos com interação) da concordância nominal de número .....	229
<b>Tabela 7</b>	– Dados comparativos: uso da concordância no SN em função da faixa etária no PB .....	231
<b>Tabela 8</b>	– Dados comparativos: uso da concordância no SN em função do gênero no PB .....	233
<b>Tabela 9</b>	– Dados comparativos: uso da concordância no SN em função do número de constituintes flexionáveis .....	234
<b>Tabela 10</b>	– Dados comparativos: uso da concordância no SN em função da saliência fônica no PB .....	236
<b>Tabela 11</b>	– Estimativas dos parâmetros do modelo (de regressão logística, modelo linear generalizado de efeitos mistos com interação) da concordância nominal de número – amostra ELiHS .....	238

## LISTA DE SIGLAS

- ABA – Associação Brasileira de Antropologia
- ADCT – Ato das Disposições Constitucionais Transitórias
- ANPOLL – Associação Nacional de Pós-Graduação e Pesquisa em Letras e Linguística
- ASCQVA – Associação Comunitária Quilombola de Volta do Angico ()
- BNCC – Base Nacional Comum Curricular
- CAAE – Certificado de Apresentação para Apreciação Ética
- CE-DOHS – *Corpus* Eletrônico de Documentos Históricos do Sertão
- CEP – Comitê de Ética em Pesquisa
- CESOL – Centros Públicos de Economia Solidária
- CNPq – Conselho Nacional de Desenvolvimento Científico e Tecnológico
- CNS – Conselho Nacional de Saúde
- CODETER – Conselho de Desenvolvimento Territorial
- CONEP – Conselho Nacional de Ética em Pesquisa
- CONSEPE – Conselho Superior de Ensino, Pesquisa e Extensão
- DCHT – Departamento de Ciências Humanas e Tecnologias
- DID – Diálogo entre informante e Documentador
- ELiHS – Estudos Linguísticos e Históricos do Sertão
- FCP – Fundação Cultural Palmares
- GT – Grupo de Trabalho
- INCRA – Instituto Nacional de Colonização e Reforma Agrária
- MDA – Ministério do Desenvolvimento Agrário
- PAB – Português Afro-brasileiro
- PB – Português Brasileiro
- PCN – Parâmetros curriculares Nacionais
- PEPP – Programa de Estudo do Português Popular de Salvador
- PHPB – Projeto para a História do Português Brasileiro
- PGEAFIN – Programa de Pós-Graduação em Estudos africanos, Povos Indígenas e Culturas Negras
- PPGEL – Programa de Pós-graduação em Estudos Linguísticos
- PRA – Plano de Recurso e Aplicação
- PROAPEX – Programa de Apoio a Projetos de Extensão
- PNDST – Programa Nacional de Desenvolvimento Sustentável dos Territórios Rurais



PTDRSS – Plano Territorial de Desenvolvimento Rural Sustentável e Solidário

SDT – Secretaria de Desenvolvimento Territorial

SEI – Superintendência de Estudos Econômicos e Sociais da Bahia

TALE – Termo de Assentimento Livre e Esclarecido

TCLE – Termo de Consentimento Livre e Esclarecido

TIC – Tecnologia de Comunicação e Informação

TII – Território de Identidade de Irecê

UEFS – Universidade Estadual de Feira de Santana

UNEB – Universidade do Estado da Bahia

UNICAMP – Universidade Estadual de Campinas

XML – Extensible Markup Language

## SUMÁRIO

<b>1</b>	<b>CONSIDERAÇÕES INICIAIS .....</b>	<b>19</b>
----------	-------------------------------------	-----------

### PARTE I – PRESSUPOSTOS TEÓRICOS

<b>2</b>	<b>PERSPECTIVA SÓCIO-HISTÓRICA DO PORTUGUÊS BRASILEIRO: CAMINHOS DO CONTATO LINGUÍSTICO .....</b>	<b>24</b>
<b>2.1</b>	<b>Contexto sócio-histórico e demográfico na formação do PB .....</b>	<b>25</b>
<b>2.2</b>	<b>Línguas africanas no Brasil .....</b>	<b>44</b>
<b>2.3</b>	<b>Formação do Português Brasileiro: algumas hipóteses e discussões .....</b>	<b>57</b>
2.3.1	Deriva Secular .....	59
2.3.2	Contato linguístico e sua aplicação ao contexto brasileiro .....	62
2.3.2.1	Conceituando pidgins e crioulos .....	63
2.3.3	Transmissão Linguística Irregular .....	69
<b>3</b>	<b>SOCIOLINGUÍSTICA: PRESSUPOSTO TEÓRICO PARA UM RETRATO DOS PROCESSOS DE VARIAÇÃO E MUDANÇA NO PORTUGUÊS BRASILEIRO .....</b>	<b>73</b>
<b>3.1</b>	<b>Linguagem e linguística: considerações gerais .....</b>	<b>74</b>
<b>3.2</b>	<b>A Sociolinguística e a sistematicidade da mudança .....</b>	<b>85</b>
3.2.1	Sociolinguística Quantitativa .....	92

### PARTE II – CONTRIBUIÇÕES PARA A PROPOSTA DE INTERIORIZAÇÃO LINGUÍSTICA

<b>4</b>	<b>EM DIREÇÃO AO SERTÃO E DO PRESENTE PARA O PASSADO: ELiHS, UM PROJETO, UM <i>CORPUS</i> E UMA PLATAFORMA <i>ONLINE</i> .....</b>	<b>97</b>
<b>4.1</b>	<b>Território de Identidade de Irecê e as comunidades rurais afro-brasileiras: uma breve caracterização sócio-histórica .....</b>	<b>99</b>
4.1.1	Caracterização sociodemográfica do Território de Identidade de Irecê .....	101
4.1.2	Dos Quilombos às Comunidades rurais afro-brasileiras: da descrição terminológica ao enquadramento histórico .....	108

4.1.3	Comunidades rurais afro-brasileiras no TII: do mapeamento cartográfico à visibilidade territorial .....	115
4.1.4	Comunidades do Território: dados sociodemográficos para uma breve caracterização identitária .....	127
<b>4.2</b>	<b>Projeto Estudos Linguísticos e Históricos do Sertão (ELiHS) .....</b>	<b>141</b>
4.2.1	Frentes de trabalho .....	144
4.2.2	Da estratificação à disponibilização das amostras: abordagens metodológicas	152
4.2.2.1	Comitê de Ética em Pesquisa .....	154
4.2.2.2	Seleção das comunidades e dos participantes .....	155
4.2.2.3	Gravação .....	163
4.2.2.4	Transcrição .....	165
4.2.3	A rede de pesquisa: parcerias importantes .....	167
<b>4.3</b>	<b>Conectividade ELiSH: armazenamento e disponibilização do acervo .....</b>	<b>173</b>
4.3.1	Concordância: descrição do suporte pedagógico para o ensino da concordância nominal .....	182
<b>5</b>	<b>A VARIACÃO NA REALIZAÇÃO DE MARCAS DE PLURAL NO SINTAGMA NOMINAL E SUA IMPORTÂNCIA PARA OS ESTUDOS SOBRE A FORMAÇÃO E A DESCRIÇÃO DO PORTUGUÊS BRASILEIRO .....</b>	<b>189</b>
<b>5.1</b>	<b>Revisitando os princípios gerais da concordância nominal: da prescrição à descrição gramatical .....</b>	<b>190</b>
5.1.1	A revisão bibliográfica: contribuições de alguns estudos .....	196
<b>5.2</b>	<b>Metodologia e instrumentos para levantamento e análise dos dados .....</b>	<b>207</b>
5.2.1	<i>Locus</i> da amostra: comunidade Volta do Angico .....	208
5.2.2	Procedimentos para a análise estatística .....	216
<b>5.3</b>	<b>Análise estatística da concordância nominal de número em amostra da comunidade afro-brasileira Volta do Angico .....</b>	<b>223</b>
5.3.1	Análises multivariadas .....	224
5.3.1.1	Análise multivariada de efeitos fixos e mistos .....	228
<b>6</b>	<b>CONSIDERAÇÕES FINAIS .....</b>	<b>240</b>
	<b>REFERÊNCIAS .....</b>	<b>246</b>

## 1 CONSIDERAÇÕES INICIAIS

*“[...] Laus Deo, que Deus seja louvado; porque o trabalho, a meu ver, será muito, de muitos e não será por pouco tempo” (Mattos e Silva, 2008a, p. 30).*

Esta tese integra a linha de pesquisa *Variação e Mudança Linguística no Português*, do Programa de Pós-graduação em Estudos Linguísticos (PPGEL) da Universidade Estadual de Feira de Santana (UEFS), objetivando contribuir com as pesquisas em torno do português afro-brasileiro (PAB), bem como da própria formação do português falado no Brasil, por meio da constituição de *corpus* oral do PAB no interior da Bahia, Território de Identidade de Irecê – Bahia. O *corpus*, representativo do português em sua vertente popular, pertence ao Projeto Estudos Linguísticos e Históricos do Sertão (doravante ELiHS), idealizado e implementando na vigência deste doutorado, sediado na Universidade do Estado da Bahia (UNEB), Departamento de Ciências Humanas e Tecnologias (DCHT), *campus* XVI – Irecê, em parceria com o *Corpus* Eletrônico de Documentos Históricos do Sertão (CE-DOHS), sediado na Universidade Estadual de Feira de Santana (UEFS).

Sendo assim, o objetivo central desta tese é a conciliação de aspectos de natureza sócio-histórica com aspectos de natureza linguística, por meio da análise variacionista laboviana da concordância nominal de número no português falado na comunidade afro-brasileira de Volta do Angico, na tentativa de que os resultados endossem as pesquisas científicas, em especial as de cunho Sociolinguístico, que estão “para além do litoral” e favorecem a discussão sobre a formação da realidade sociolinguística brasileira.

Dados sobre uso variável da concordância nominal no PB têm sido recorrentes nas últimas décadas, em monografias, dissertações e teses, as quais favorecem um panorama sistemático dessa variação no PB; no entanto, o que potencializa a abordagem nesta tese é o fato de proporcionar a discussão sobre a formação da realidade sociolinguística brasileira, através da análise de um *corpus* inédito, constituído no âmbito desta pesquisa, que se desdobrou em um Projeto de Pesquisa – o ELiHS.

Assim, o estudo da variação da concordância nominal de número, fenômeno gramatical objeto de análise desta tese, tende a buscar indícios de que marcas linguísticas presentes em amostras de falas sincrônicas podem ser explicadas através da lupa socio-história, em que normas linguísticas foram demarcadas, consequência da colonização exploratória firmada no Brasil.

Neste contexto, iniciamos nosso primeiro capítulo desta tese, intitulado *Perspectiva sócio-histórica do português brasileiro: caminhos do contato linguístico*, em que comungamos

com a hipótese da bipolarização do PB, marcada, fortemente, pela presença das línguas africanas, cujos falantes foram forçados a adquirirem o português como segunda língua, em um contexto que assumimos aqui de *Transmissão Linguística Irregular*. Arrolaremos, também, algumas hipóteses, levantadas na literatura sobre a formação do PB, bem como as condições sócio-histórico-demográficas que, possivelmente, favoreceram uma aquisição ‘imperfeita’/defectiva do PB, língua herdada pelos afro-brasileiros, condicionando o que chamaremos de bipolarização linguística. Acreditamos, preliminarmente, que tais marcas de uma aquisição defensiva, de forma mais incisiva por parte dos afrodescendentes, tenham sido incorporadas em seu contexto linguístico e sistematizadas enquanto língua de um povo. Assim, levanta-se a seguinte questão: ao serem comparadas com outros estudos já realizados sobre a concordância nominal, essas marcas variáveis estariam mais presentes na oralidade de povos que, de alguma forma, se afastaram dos centros urbanos e foram excluídos dos processos de escolarização, assim como em comunidades Quilombolas, hoje representadas por seus descendentes, residentes em Comunidades Remanescentes de Quilombolas? Portanto, partiremos da hipótese de que um contingente expressivo adquiriu uma gramática em que configuravam sintagmas nominais sem a presença da marca de plural canônica, a qual acreditamos estar diretamente relacionada aos afro-brasileiros, principalmente os que ainda vivem com comunidade rurais.

Dissimilaridades existentes linguisticamente entre norma popular e culta, rural e urbana têm impulsionado estudiosos a compreenderem quais foram os principais vetores intensificadores nos processos de variações e mudanças linguísticas. Portanto, nesta tese, assumiremos a língua como um produto social, através dos pressupostos da Sociolinguística Quantitativa. Assim, no capítulo intitulado *Sociolinguística: pressuposto teórico para um retrato dos processos de variação e mudança no Português Brasileiro*, traremos um percurso dos estudos sobre a variação e, principalmente, sobre a mudança linguística, abordando desde concepções formalistas até sócio-histórico-culturais, em que teremos a Sociolinguística como eixo teórico norteador desta pesquisa.

Os dados demográficos, da historiografia e da mobilidade populacional do Brasil no período colonial, apresentado na *Seção 2*, se somarão a um conjunto de fatores para a compreensão das motivações sócio-históricas que deram origem ao português popular brasileiro, permitindo entender, de forma mais específica, o contexto em que as comunidades rurais afro-brasileiras, que compõem nossa amostra, estão inseridas. Dessa forma, no capítulo intitulado *Em direção ao sertão e do presente para o passado: ELiHS, um projeto, um corpus e uma plataforma online*, apresentaremos dados de demografia histórica, de mobilidade

sociocultural e geográfica, além de aspectos político-econômicos como fatores relevantes na recomposição de quadros sociolinguísticos do Território de Identidade de Irecê. Atrrelados às discussões, apresentamos a caracterização e o mapeamento das mesmas, fazendo-as presentes no mapa da sociolinguística brasileira. Soma-se a este objetivo o de apresentar informações que subsidiem a análise sociolinguística da concordância nominal de número na variedade falada na comunidade afro-brasileira Volta do Angico (BA). Ainda neste capítulo, descreveremos os critérios metodológicos que subsidiaram a implementação do Projeto ELiHS e apresentaremos, também, a plataforma *online* do Projeto, criada através de uma rede colaborativa, em que se busca armazenar e popularizar dados orais sincrônicos, representativos do português rural afro-brasileiro falado no Território de Identidade de Irecê (TII) – Ba, possibilitando, por meio da criação de plataforma *online* gratuita, a qualquer interessado ver, ouvir, repetir os usos linguísticos das comunidades representativas do Território, subsidiando novas pesquisas científicas – dentro e fora da UNEB.

No último capítulo, e não menos importante, faremos uma explanação acerca do tema morfossintático escolhido para subsidiar a análise sobre a formação da realidade sociolinguística brasileira: a concordância de número. Tentaremos realizar um apanhado geral das pesquisas já realizadas no Brasil, para, por fim, apresentarmos os resultados acerca da variação analisada, testando o comportamento das variantes em relação aos grupos de fatores, comparando os resultados obtidos com os alcançados em outras pesquisas, em especial da pesquisa realizada por Andrade (2003), na comunidade afro-brasileira de Helvécia. Sendo assim, no capítulo intitulado *A variação na realização de marcas de plural no sintagma nominal e sua importância para os estudos sobre a formação e a descrição do português brasileiro*, analisaremos o fenômeno na amostra coletada, buscando fornecer subsídios para a confirmação da hipótese de que a concordância nominal de número é uma regra variável, condicionada por fatores linguísticos e sociais, e contribuindo para uma proposta de mapeamento da variabilidade do fenômeno, assim como para a sua interiorização, nas vias do processo de variação e mudança.

O delinear da proposta, descrita ainda brevemente nestas *Considerações Iniciais*, atribui à tese uma relevância pontual, não apenas no âmbito científico, mas principalmente no âmbito Territorial e, de forma mais específica, Departamental – ao curso de Letras do DCHT – *campus XVI*. Certamente, a defesa desta tese representará apenas um marco temporal, visto que seus frutos continuarão florescendo (como propõe a imagem visual do Projeto), diante da implementação do primeiro Projeto de Pesquisa em Linguística do curso de Letras da UNEB – *Campus XVI/Irecê*.

## **PARTE I – PRESSUPOSTOS TEÓRICOS**

## 2 PERSPECTIVA SÓCIO-HISTÓRICA DO PORTUGUÊS BRASILEIRO: CAMINHOS DO CONTATO LINGUÍSTICO

“[...] O português são dois; o outro, mistério.”  
Carlos Drummond de Andrade

Como ponto de partida, tomaremos a proposta de Mattos e Silva (2001, p. 298-299), visto que esta tese, dentre seus objetivos norteadores, busca contribuir para compreensão em torno do processo de formação do Português Brasileiro, especificamente o Português em sua vertente popular, a partir da constituição de amostras de falas em comunidades afro-brasileiras, situadas no interior da Bahia, na tentativa de conciliar aspectos de natureza sócio-histórica com aspectos de natureza linguística. A abordagem será guiada por Mattos e Silva (2001) e Lucchesi (1994, 2001 e 2006) – que também advoga em prol de uma formação histórica do PB –, a partir de duas macro origens distintas, que se resumem em termos de uma polarização<sup>2</sup>: de um lado, uma *norma culta* derivada dos padrões linguísticos da elite da Colônia e do Império; e, de outro, as *variedades populares* do PB, marcadas por um conjunto de mudanças estruturais induzidas pelo contato entre línguas, através do processo de transmissão linguística irregular (cf. Lucchesi, 2003, 2008).

Também apresentaremos as principais propostas de periodização<sup>3</sup> para a história da língua portuguesa no Brasil, em especial a de Lucchesi (2017), a qual é inovadora por levar em consideração um conjunto mais amplo de línguas presentes na formação da sociedade brasileira, além de tomar como base os princípios do Paradigma Variacionista, teoria que subsidia nossas análises, buscando correlacionar os processos linguísticos aos processos sociais.

Nesse contexto, defenderemos que a expressão “variedades populares do português do Brasil” designa a fala representativa de comunidades de baixa ou nenhuma escolarização, em relação ao que Faraco (2008) e Lucchesi (2002) determinam de norma culta brasileira. Assim, acreditamos que o fato de as amostras de fala serem de comunidade afro-brasileira, representativas do português popular, colabora para o entendimento das suas diferenças em

---

<sup>2</sup> Lucchesi (1994; 2001) concebe o português do Brasil como um sistema não apenas heterogêneo e variável, mas plural, um diassistema formado por dois subsistemas, por sua vez, igualmente heterogêneos e variáveis, definidos como “normas”. Distingue-se, fortemente, de Silva Neto (1963 [1951] p. 234-235) sobre a unidade e o conservadorismo no português do Brasil e, sobretudo, da visão de uma suposta hierarquia de culturas.

<sup>3</sup> Silva Neto (1963 [1951]), Marlos Pessoa (2003), Tania Lobo (2003), Volker Noll (2008) e Dante Lucchesi (2017).



relação à fala de camadas urbanas e mais escolarizadas (Baxter, 1997; Baxter; Lucchesi, 1997; Lucchesi, 2001, 2002, 2009), principalmente no que tange ao fenômeno aqui estudado da variação na concordância nominal de número. Ainda importante destacar que a investigação em torno do português popular em comunidade afro-brasileira, como propõe esta pesquisa, faz-se urgente por acreditarmos que seu antecedente histórico teria sido, segundo Mattos e Silva (2001), o português geral brasileiro, adquirido em situações de aquisição imperfeita, difundido pelo Brasil através dos africanos e afrodescendentes, “[...] principais difusores da língua portuguesa no Brasil e os principais formatadores do português brasileiro em sua variante social majoritária — o português popular brasileiro” (Lobo, 2015, p. 69).

A discussão sobre a formação e a difusão do português falado no Brasil nos permite inferir, parafraseando Drummond, que o português são vários, e que, como define Mattos e Silva (2004a, p. 14), vai de “[...] um multilinguismo generalizado a um multilinguismo localizado”. Portanto, cabe à Linguística, em especial à Sociolinguística, desvendar seus mistérios, fazendo “bom uso dos maus dados”, numa perspectiva do presente para o passado, como propõe esta tese.

## **2.1 Contexto sócio-histórico e demográfico na formação do PB**

No contexto de formação do português brasileiro, buscaremos dar lugar e voz aos milhares de indígenas, que já viviam em terras brasileiras, com culturas e línguas autônomas, como bem descreve o professor Ayron Rodrigues (1993), e também aos negros africanos, que por muito tempo tiveram suas línguas, vozes e histórias ocultadas na formação da nossa língua, conforme aponta Petter (2007). É através desse viés que buscaremos romper com o ideal eurocêntrico, impregnado nas abordagens tradicionais da história linguística do Brasil, que põe à margem toda a história criada/vivida pelos autóctones, africanos e afrodescendentes, e colaborar, mesmo que de forma descritiva, para uma aproximação da compreensão do português brasileiro – tão sonhada pela Rosa Virgínia Mattos e Silva, que, como homenageia Lobo (2015, p. 69), “[...] representa, na segunda metade do século XX, um pensamento vivo e instigante sobre a história social linguística do Brasil”.

Aqui, tomaremos como base reflexiva as 05 (cinco) proposições apontadas por Lobo (2015, p. 71), a partir de uma releitura sintética e objetiva dos 10 (dez) textos<sup>4</sup> publicados pela professora Mattos e Silva, a saber:

- A história linguística do Brasil não se restringe à história da língua portuguesa no Brasil, nem à história do português brasileiro;
- O português brasileiro emerge em contexto multilíngue: o contato linguístico é, pois, elemento constitutivo da sua formação;
- Na cena linguística do Brasil colonial, destacam-se três atores principais: o português europeu, as línguas gerais indígenas e o português geral brasileiro;
- Africanos e afrodescendentes foram os principais difusores da língua portuguesa no Brasil e os principais formadores do português brasileiro em sua variante social majoritária — o português popular brasileiro;
- O passado sócio-histórico-linguístico do Brasil deverá ser interpretado para a compreensão do português brasileiro «heterogêneo e variável, plural e polarizado» da atualidade.

Assim, partimos da ideia de que descrever, ainda que brevemente, o passado da sociedade brasileira colaborará, como afirma Mattos e Silva (2001), para uma aproximação menos generalizante da história social linguística do Brasil e, conseqüentemente, para a história do português brasileiro. Nesse sentido, Silva Neto (1986, p. 14, grifos do autor) já ressaltava que

a matéria de história da língua portuguesa no Brasil há de investigar-se na etnografia e na evolução histórico-social do povo brasileiro.  
Esse assunto encontra símiles na história externa do português asiático e africano, do inglês norte-americano, do espanhol argentino, chileno etc.  
Mas é preciso frisar a muito maior complexibilidade do caso brasileiro, em vista das particularidades de nossa formação *étnico-social*.

Esse contexto de particularidade pode começar a ser narrado com o processo de colonização, em que, a partir da década de 30 do século XVI, acontecem os principais contatos dos portugueses com as línguas indígenas no Brasil, quando o número de colonos ainda era pequeno, e esses dependiam dos índios para conhecer e conseguir sobreviver no Brasil.

---

<sup>4</sup> Textos presentes na obra *Ensaio para uma sócio-história do Português brasileiro* (2004a).

Importante a ressalva para o fato de os colonizadores, falantes de língua portuguesa, que foram trazidos ao nosso continente, virem das mais diversas regiões de Portugal, do campo ou da cidade, configurando, portanto, um contingente já permeado pelo contexto da heterogeneidade linguística, mas, pertencendo ao mesmo domínio linguístico, os falantes eram capazes de se entender uns aos outros (Silva Neto, 1986, p. 42).

Esse contingente de falantes de língua portuguesa, em contexto heterogêneo, ao chegarem ao Brasil, encontraram a terra povoada por índios que, do ponto de vista linguístico, conviviam em um contexto multilíngue, comunicando-se através de uma espécie de língua franca – Rodrigues (1986) faz uma projeção inicial de 1.078 línguas indígenas faladas no território brasileiro, sendo o quantitativo posteriormente atualizado para 1.175 línguas (Rodrigues, 2006). É um quantitativo expressivo quando pensamos ser este o período inicial da implantação da língua portuguesa no país.

Ciente da complexidade do português falado no Brasil, Venâncio (2007) discute, de forma reflexiva, sobre a inviabilidade de se reconhecer a língua portuguesa como idioma oficial do Brasil, visto que os fatos e história narram que, em 1500, o atual território brasileiro era povoado por povos indígenas, falantes de idiomas distintos, além disso, ainda nesse contexto, ocorreu a vinda, forçada, de milhões de africanos, em “[...] 90% dos casos provenientes da África Atlântica, compreendida *grosso modo* entre Angola e Senegâmbia” (Venâncio, 2007, p. 359, grifos do autor). Assim, além das línguas gerais e das inúmeras línguas indígenas, o português concorreu com as línguas dos africanos de diferentes grupos étnicos, que desde o início foram trazidos ao Brasil.

Como assinala Mattos e Silva (2004), os negros foram fundamentais nesse cenário sócio-histórico, uma vez que, em substituição ao trabalho indígena, foi crescente a imigração forçada dos negros africanos para o Brasil. Esse processo de tráfico de africanos, em que eram submetidos ao trabalho escravizado, inicia-se nos primeiros anos de colonização, oficialmente em 1549, por Dom João III, e perdura por um longo período. Segundo Américo Venâncio Machado Filho (2008), esses milhões de negros africanos pertenciam, principalmente, a dois troncos linguísticos: o afro-asiático e o congo-cordofaniano. O autor afirma que o primeiro representa uma minoria e se fez presente sobretudo na região do Nordeste, especialmente na Bahia; o segundo pertence à maioria dos escravos que chegaram ao Brasil, e integrava dois principais ramos de uma única família linguística – Níger-Congo.

Castro (2009, p. 179) apresenta o seguinte quadro de línguas subsaarianas documentadas no Brasil durante o período da escravidão, de acordo com a proposta de Joseph Greenberg, em 1966.

**Quadro 1** – Línguas subsaarianas

FAMÍLIA	GRUPO LINGUÍSTICO	LÍNGUA	SÉCULO DE REGISTRO	LOCALIZAÇÃO	
Níger-congo	Banto	Congo-Angola	XVII ao XIX	Salvador (Bahia)	
	Oeste-africanas	KWA	Ewe-fon ou mina-jeje	XVIII	Zona de mineração – Minas Gerais
			Nagô-iorubá	XIX	Centro urbano – Salvador (Bahia)
			Tapa ou nupe	XIX	Centro urbano – Salvador (Bahia) – minoritária
		Atlântico ocidental	Fulani	XIX	Centro urbano – Salvador (Bahia) – minoritária
		Gur ou Voltáico	Grunce	XIX	Centro urbano – Salvador (Bahia) – minoritária
Afro-asiática	Afro-asiático	Hauçá	XIX	Centro urbano – Salvador (Bahia) – minoritária	

**Fonte:** Castro (2009, p. 179).

O *Quadro 1* nos permite afirmar que, apesar da diversidade de línguas transplantadas para o Brasil, é possível traçar uma origem comum, uma vez que as línguas, com exceção da Hauçá, da família afro-asiática, pertencem à única família linguística Níger-congo, o que nos permite afirmar que as línguas africanas vindas ao Brasil eram línguas aparentadas.

Mattos e Silva (2004) e outros autores também apresentam argumentos veementes que comprovam que não só a língua portuguesa esteve presente no processo de formação da sociedade brasileira, mas também as indígenas e africanas. Porém afirmam que os dados demográficos, apresentados em seus trabalhos, representam apenas estimativas demográficas, pois poucos registros se têm para as primeiras décadas da colonização no Brasil. Na *Tabela 1*, podem-se observar valores relativos à população indígena, em comparação à população europeia e africana:

**Tabela 1** – Estimativas de população no Brasil

PERÍODO	POPULAÇÃO INDÍGENA	POPULAÇÃO EUROPEIA E AFRICANA	POPULAÇÃO TOTAL	% DE POPULAÇÃO INDÍGENA
c. 1500	4.000.000	-	4.000.000	100,0
c. 1822	800.000	3.596.132	4.396.132	18,1
c. 1889	215.000	13.733.915	13.948.915	1,5

Fonte: Venâncio (2007, p. 361).

As estimativas demonstram que, nos primeiros anos de colonização, a população indígena era majoritária, mas aos poucos esse cenário sofre alterações, com a inserção da população europeia e o início do tráfico negreiro em terras brasileiras. Em 1889, pode-se visualizar uma pequena parcela da população indígena, que corresponde a apenas 1,5% da população geral, o que traduz o pronunciamento enfático de Mattos e Silva (2004, p. 14, grifo nosso) ao afirmar que no processo de colonização do Brasil “[...] se inicia a trajetória dizimadora dos índios brasileiros e de suas línguas, *percurso etnocida e glotocida* [...]”. Essas poucas e sábias palavras conseguem retratar todos os números apresentados por Venâncio (2007).

Com isso, pode-se considerar que a historiografia sobre a colonização brasileira demonstra que não foi apenas o elemento português que entrou na demografia do Brasil, mas também os autóctones e os africanos, bem como outros europeus que imigraram para o Brasil na condição de trabalhadores assalariados desde a primeira metade do século XIX. Para maior precisão, faz-se necessário apresentar estimativas que demonstrem a imigração para o Brasil entre 1651 e 1800.

**Tabela 2** – Estimativas de imigração para o Brasil

PERÍODO	PORTUGUESES	AFRICANOS	TOTAL DE IMIGRANTES
1651-1700	100.000	360.000	460.000
1701-1800	800.000	1.700.300	2.500.300

Fonte: Venâncio (2007, p. 361).

Sabe-se que, na medida em que imigrantes europeus chegavam ao Brasil e africanos eram introduzidos através do tráfico, a quantidade de indígenas ia sendo reduzida, porém,

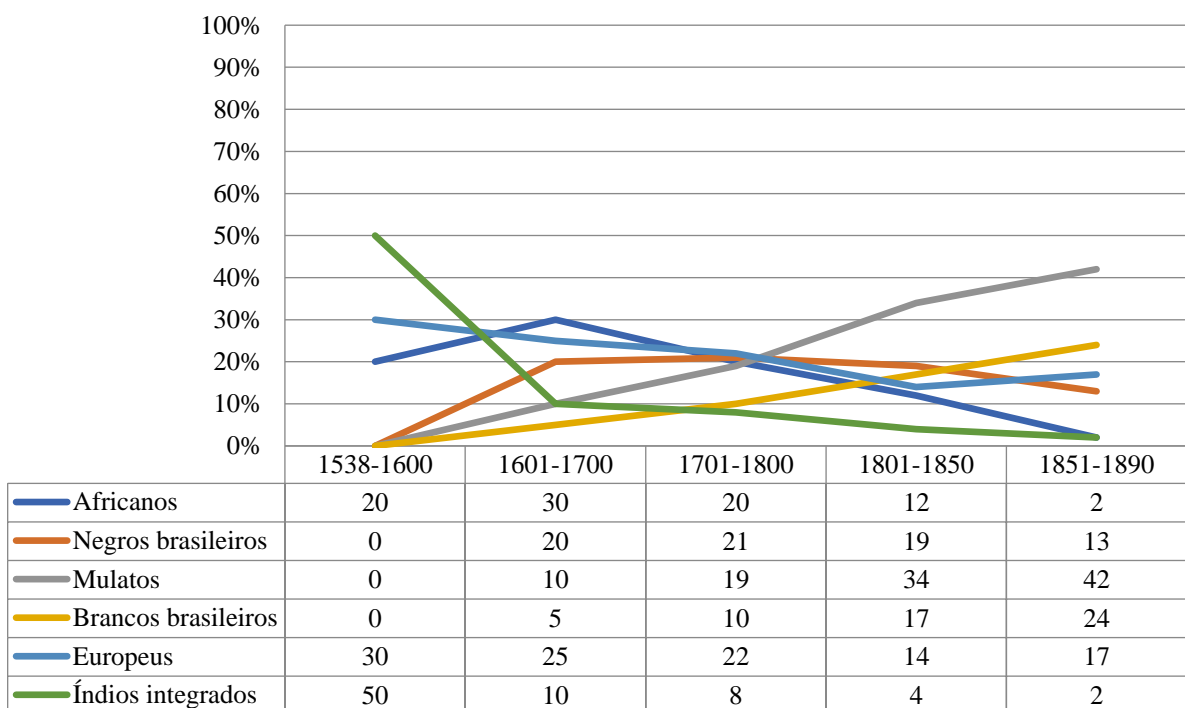
durante um longo período, apresentado na *Tabela 2*, esses povos com etnias, línguas e culturas diversas conviviam em um mesmo território, configurando, ou melhor, intensificando um cenário de multilinguismo, uma vez que, como afirma Lucchesi (2009), o português brasileiro apresentava nesse contexto não apenas um caráter heterogêneo e variável, mas também plural e polarizado, visto que

[...] o processo do encontro entre a língua portuguesa, língua de dominação, com muitas línguas autóctones e as diversas línguas aqui chegadas, primeiro as africanas, depois as línguas de imigrantes, que tornaram esta área americana, multilíngue de origem, ainda mais complexa linguisticamente (Mattos e Silva, 2004, p. 11).

Os dados/números apresentados nas *Tabelas 1 e 2* são essenciais, pois “informações históricas desse tipo são preciosas para a reconstituição histórica do passado linguístico do PB, [...] já que não se pode negar que a história das línguas passa necessariamente pela história demográfica de seus falantes” (Mattos e Silva, 2004, p. 17). Silva Neto (1986, p. 19, grifos do autor), já nos dizia que “[...] o ponto fundamental desta centenária querela de *língua brasileira* ou *dialeto brasileiro*, residiu no erro palmar de encarar-se o nosso português como *um bloco, uma massa uniforme*”.

Com dados mais detalhados sobre a demografia histórica do Brasil, Mussa (1991, p. 163) apresenta as porcentagens étnicas presentes nas diversas fases de colonização:

**Gráfico 1** – Demografia histórica do Brasil



**Fonte:** Adaptado de Mussa (1991, p. 163).

Considerando os dados demográficos referentes ao período da colonização do território brasileiro, é possível considerarmos que nos séculos XVII e XVIII já se observa um cenário pluriétnico, com população “equilibrada” entre europeus, africanos, indígenas e a primeira geração desses povos (mulatos, negros brasileiros e brancos brasileiros). Comparando as primeiras décadas de colonização com as últimas décadas do século XIX, fica visível um drástico declínio da população indígena (de 50% para 2%), dando espaço ao aumento de mulatos (de 0% a 42%), e de brancos brasileiros (de 0% a 24%). Mattos e Silva (2004a) afirma que o aumento populacional de europeus e brancos brasileiros, na segunda metade do século XIX, é reflexo da extinção do tráfico, coincidindo com a chegada de imigrantes da Europa e da Ásia. Em contraponto, tem-se uma ascensão da população africana (de 20% para 30% – século XVI e XVII) e, em seguida, com a extinção do tráfico negreiro, um declínio progressivo, chegando ao final do século XIX com apenas 2% da população geral.

É a partir dos dados apresentados por Mussa (1991) que Mattos e Silva (2004a) levanta argumentos que tentam explicar a generalizada difusão do português popular brasileiro, recaindo nos encontros interlinguísticos entre indígenas, africanos e afrodescendentes. Nesse sentido, Mattos e Silva (2004a) reúne alguns fatores que sustentam sua hipótese, a saber:

- (i) a presença massiva de negros e afrodescendentes ao longo de toda a colonização;
- (ii) a atuação constante dos escravos nas grandes frentes da economia colonial;
- (iii) a mobilidade geográfica dos escravos;
- (iv) o significado social e linguístico dos espaços ilegítimos da escravidão.

Diante das pesquisas realizadas no campo da etnolinguística, já se pôde registrar as muitas línguas indígenas (Rodrigues, 1986) e africanas (Castro, 2001) que se fizeram presentes nos primeiros séculos. Porém esses povos, índios e africanos, representam, no século XIX, uma pequena parcela da população (*cf.* Gráfico 1), o que nos faz levantar as seguintes questões: a) se já se tem certeza do cenário plurilíngue montado em terras brasileiras, como se caracterizou o processo de intercâmbio entre índios, africanos e europeus?; b) o que restou desses povos, das suas respectivas línguas, no português falado hoje, no século XXI, no Brasil?. Essas são algumas das perguntas que nos norteiam, quando se pensa no processo de constituição do Português Brasileiro.

Essas questões ecoam na indagação levantada por Teyssier (1997, p. 97) no livro *História da Língua Portuguesa*, “Como explicar as particularidades do português do Brasil?”; porém sabemos que a resposta a essa indagação constitui um dos problemas cruciais na busca

de resultados empíricos para se alcançar o conhecimento referente à efetiva formação do português brasileiro.

No que tange ao processo de colonização e à difusão do português no Brasil, Mattos e Silva (2004, p. 14) conta-nos, através da narrativa de Pero Vaz de Caminha, que quatro portugueses, dois degredados e dois grumetes que fugiram da frota, ficaram em solo brasileiro, e foram responsáveis por semear o português no Brasil, mas, antes, estes foram “obrigados” a aprender a língua dos índios, para, só então, ensinar-lhes o português. Negrão e Viotti (2012, p. 14) afirmam que, até aproximadamente 1530, a Coroa Portuguesa enviava navios para o Brasil com dois objetivos: guardar a costa, protegendo a nova colônia da invasão de piratas; e extrair o tão valioso pau-brasil. Nesses navios, vinham basicamente homens, a maioria de origem portuguesa, mas não todos – havia europeus de outras nacionalidades e línguas.

Negrão e Viotti (2012) refletem sobre as pessoas envolvidas na expansão imperial europeia e deduzem que esses homens certamente deviam estar acostumados a situações de contato linguístico, e provavelmente já deviam ter desenvolvido eficientes estratégias de comunicação com falantes de línguas diferentes. Sendo assim, o português que chegou ao Brasil

não se tratava de um português único, mas de uma multiplicidade de dialetos portugueses; não se tratava de um português “puro”, mas de um português que já carregava marcas do contato com uma variedade de línguas africanas e com as línguas de outros europeus que participavam da expansão mercantilista. Uma vez em território brasileiro, as relações entre os europeus e as diversas populações indígenas que viviam na costa brasileira certamente devem ter causado o *surgimento de um sistema de comunicação funcional que permitisse a interação entre eles* (Negrão; Viotti, 2012, p. 15, grifo nosso).

O que as autoras definem como “sistema de comunicação funcional”, pode ser entendido, dentro do contexto de colonização brasileira, como a *língua geral*<sup>5</sup>. Essa língua geral nasceu quando os primeiros povoadores portugueses, no início do século XVI, vieram para o Brasil e entraram em contato com as tribos indígenas que habitavam a costa e, por falarem línguas muito aparentadas do tronco tupi, eram capazes de se comunicar entre si utilizando uma espécie de língua franca, que viria a ser denominada *língua geral da costa brasileira* – ou simplesmente *língua geral* –, a qual foi o instrumento de comunicação adotado pelos portugueses para integrar a força de trabalho indígena, inicialmente na extração do pau-brasil e posteriormente no cultivo da cana-de-açúcar, do tabaco e do algodão (Lucchesi, 2009, p. 43).

---

<sup>5</sup> A denominação *língua geral* recobre uma grande diversidade, conforme alerta Mattos e Silva (2001, p. 286): “só podemos idealizar essa língua geral como heterogênea desde o século XVI, heterogeneidade que se tornará mais complexa ao longo da diacronia da colonização [...]. Para a compreensão do que se chama genericamente língua geral, é fundamental o estudo de Aryon Rodrigues (1986: 99-109), em que, com precisão, apresenta uma caracterização diatópica e diacrônica das línguas gerais, já no plural, a paulista e a amazônica”.



Isso pode ser ratificado no trecho em que Serafim da Silva Neto (1986) declara que a *língua geral* foi uma “[...] língua comumente falada na costa, que chegou até nós com o nome de ‘geral’. Ela era precisa para comunicação”.

Assim, entendemos, nos termos de Rodrigues (1993, p. 96), que as línguas gerais

[...] não se desenvolveram como pidgins nem como crioulos, mas são continuações de línguas indígenas que passaram a ser faladas pelos mestiços de homens europeus e mulheres índias. No processo de constituição das línguas gerais deve ter havido uma sucessão de fases, incluindo um período em que os homens europeus eram poucos e viviam nas comunidades indígenas ou próximo a elas, usando eles mesmos a língua indígena como segunda língua, e seus filhos mestiços (mamelucos ou caboclos) se criavam imersos na língua indígena praticamente inalterada; outro período em que os elementos europeus e mestiços passaram a ser mais numerosos, ou porque passaram a formar comunidades mais separadas das comunidades indígenas, ou porque os indígenas foram sendo progressivamente reduzidos em suas próprias comunidades; outro período, ainda, em que, afastados ou eliminados os índios, os mestiços passaram a ser os únicos falantes nativos da língua original e os transmissores dela a outros índios e outros europeus.

Ainda nesse contexto de língua geral, temos a contribuição dos jesuítas que, por considerarem que a variedade de línguas impedia a conversão, procuraram aprender o tupi, um tupi simplificado, despojado de seus traços fonológicos e gramaticais mais típicos – sendo essa língua estudada, fixada em catecismos, dicionários e gramáticas e institucionalizada como língua de contato entre colonizadores e índios. Nesse sentido, Mattos e Silva (2004, p. 14-15) afirma que

o processo colonizador e evangelizador dos séculos XVI e XVII teve de utilizar, como instrumento fundamental para a dominação, línguas indígenas brasileiras. Embora homogeneizadora da atividade catequética – construiu um ‘tupi jesuítico’ –, a língua geral da costa, de base tupi, chegou a ser um risco para a hegemonia do português no Brasil, juntamente com outras línguas gerais indígenas que foram veículos de intercomunicação entre brancos, negros e índios não só no litoral brasileiro, mas nas entradas paulistas [...].

O que Mattos e Silva (2004) define como um risco à hegemonia do português no Brasil diz respeito à amplitude de uso da língua geral entre os habitantes das terras brasileiras, uma vez que foi utilizada como principal veículo de comunicação entre brancos, negros e índios. Diante desse contexto, em 1757, a Coroa proibiu que a língua portuguesa coexistisse com outras línguas, instaurando um decreto, pelo Marquês de Pombal, que fixava o português como língua oficial brasileira, proibindo o uso da língua geral. Dessa forma,

[...] se define uma política linguística e cultural que fez mudar de rumo a trajetória que poderia ter levado o Brasil a ser uma nação de língua majoritária indígena, já que os dados históricos informam que uma língua geral de base indígena ultrapassara de muito as reduções jesuíticas e se estabelecia como língua familiar no Brasil eminentemente rural de então (Mattos e Silva, 2004, p. 20-21).

De fato, a partir da segunda metade do século XVIII, com o decreto do Marquês de Pombal, proibindo o uso da língua geral nas escolas, e com a expulsão dos jesuítas, em 1754, a língua geral perdeu seus principais protetores, verificando-se, então, o predomínio da língua portuguesa.

Cabe relatar, portanto, que até o estabelecimento da política linguístico-cultural de Pombal, o agente escolarizador generalizado no Brasil foi a Companhia de Jesus. Mattos e Silva (2004) discorre sobre a educação na Capitania da Bahia com base na Carta VIII, do professor Luís dos Santos Vilhena, que descreve como se processava o ensino na Companhia de Jesus:

No tempo em que existiam os jesuítas, incumbido então de todas as escolas menores, tanto em Portugal, como por todos os seus domínios, havia nos gerais do Colégio desta cidade sete classes em que se instruía a mocidade, não só da capital, como de todas as vilas... na primeira das sete mencionadas classes se ensinava gramática portuguesa, desta passavam os meninos a aprender, na segunda os primeiros rudimentos da língua latina, estudavam a sintaxe e a sílaba na terceira classe, da qual passavam para a quarta, onde aprendiam a construção da mesma língua, a retórica, tal como se ensinava; na sexta filosofia e na sétima teologia moral (Vilhena, 1969, p. 273-273 *apud* Mattos e Silva, 2004, p. 39-40).

Mattos e Silva (2004, p. 40) continua comentando que, através da Carta VIII de Vilhena (1969), é possível inferir que, de fato, a língua portuguesa não era objetivo prioritário no currículo das “escolas menores” dos jesuítas. Portanto, com o decreto Pombalino, criam-se “as escolas das primeiras letras”. A autora ainda relata que, através da Carta V de Vilhena (1969), temos um *flash* do iletramento, então generalizado: “[...] são uns mulatos ou negros tão estúpidos, que eu não conheci ainda um que soubesse ler, ou escrever o seu nome, e se algum branco exercita a arte, nada difere daqueles, quanto à instrução” (Vilhena, 1969, p. 189 *apud* Mattos e Silva, 2004, p. 40).

Com a interrupção do trabalho nas escolas jesuíticas no Brasil, o esforço de escolarização da população local – indígena, mestiça, ou descendente de portugueses – não é retomado senão mais de uma década depois, de forma bastante intermitente e irregular. Ao se tratar dessa temática no Brasil, as fontes ainda são escassas, mas podemos citar o consistente trabalho de Houaiss (1985), ao apresentar dados que demonstram que, ainda no século XIX, apenas cerca de 0,5% da população era letrada, ou seja, em uma população com 3,5 a 4 milhões de habitantes, esse número correspondia a aproximadamente 18 a 20 mil letrados, os quais pertenciam à classe dirigente e dominante, ou eram “assessores” intelectuais dessa classe – como escritores, agregados, escrivães, escriturários, correspondentes, comerciantes e clero (Houaiss, 1985, p. 141).

Ao analisar o panorama da escolarização no Brasil, Mattos e Silva (2004) conclui que os dados sobre a escolarização permitem-nos entrever que são as variantes orais dialetais espaciais e verticais que dominam o português brasileiro. Já em relação à normatização da língua, desencadeada no século XIX, a autora acredita que não houve uma generalização para o território brasileiro, uma vez que, por ser de caráter lusitanizante, restringiu-se apenas à elite brasileira.

Porém, mesmo com a implantação dos agentes promotores da normatização linguística – a escola –, não temos a exclusão das centenas de línguas que vigoravam no país. Nesse sentido, Lucchesi (2001, p. 102) afirma que

[...] a língua portuguesa tinha que lutar para se impor, em primeiro lugar, contra às variedades da língua franca de base tupi falada na costa brasileira, chamada *língua geral*. Por outro lado, a língua portuguesa era adquirida nas situações as mais precárias pelos escravos, que muitas vezes preferiam se comunicar entre si, usando uma língua franca africana. Sob essas ásperas condições, a língua portuguesa foi se disseminando entre a população pobre, de origem predominantemente indígena e africana, nos três primeiros séculos da história do Brasil.

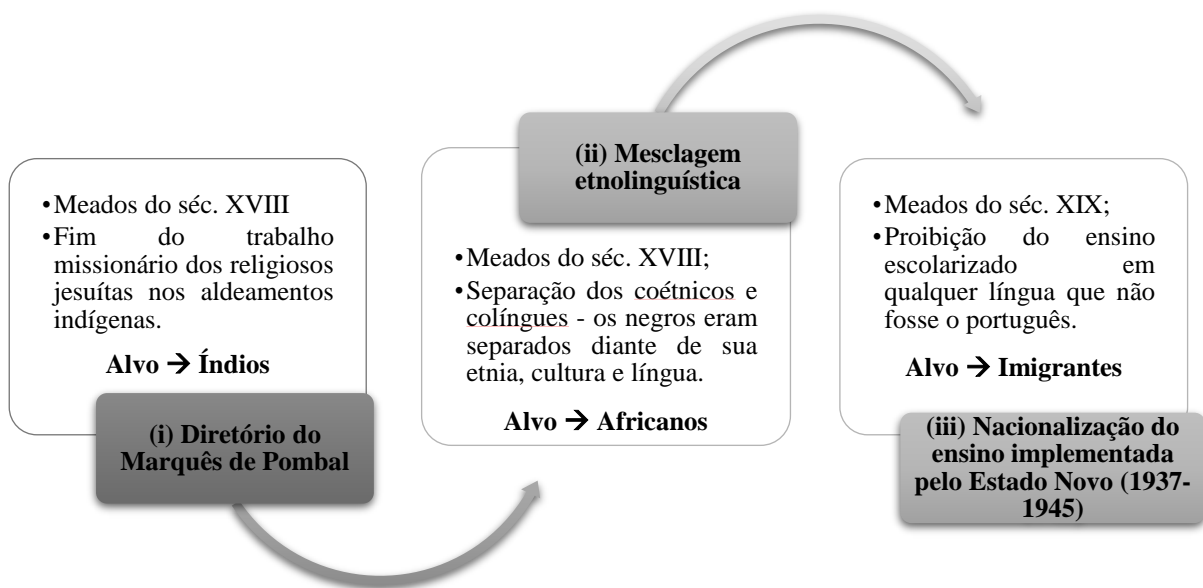
Sendo assim, Serafim da Silva Neto (1986, p. 36) afirma que “[...] há muitas semelhanças entre o português dos índios e o português dos negros. Isso é, aliás, bem natural, pois tanto o índio como o negro, em atraso de civilização, aprenderam o português como língua de emergência, obrigados pela necessidade”.

Até porque, o autor continua comentando, a educação não compete exclusivamente ao professor, mas encontra sólido apoio no convívio com as pessoas que nos rodeiam, nas coisas que vemos e até nas maneiras que nós observamos nos outros (Silva Neto, 1986, p. 40). O contato entre portugueses, índios e africanos, designado por Silva Neto (1986, p. 40) como co-educação, gerou influências mútuas, no que tange à fusão das culturas, dando origem a uma sociedade híbrida, de caráter diverso da europeia. Ao tratar do vocabulário dos povoadores portugueses, acredita-se que esse “restringia-se por força, porque no Brasil não se encontravam dezenas e dezenas de coisas ou fatos do Reino. Mas, por outro lado, enriquecia-se lhes a expressão com o adotar de numerosíssimos vocábulos designativos de coisas e fatos americanos” (Silva Neto, 1986, p. 40). Diante de tal contexto, o autor questiona: enriqueceu-se no Brasil a linguagem corrente portuguesa ou, pelo contrário, diminuíram-lhe os recursos? Busca-se até hoje a resposta para esse questionamento.

A descrição, ainda que horizontalizada, traçada, nesta seção, sobre os primeiros séculos de colonização, demarca que o português, no Brasil, foi implantando sob, como afirma Silva Neto (1986), um contexto de particularidades referentes à formação *étnico-social* que condicionam hoje um distanciamento em relação ao português lusitano (Silva Neto, 1986, p.

28). Uma condição que embute no pesquisador linguístico a inquietação investigativa sobre o processo de implantação do português em terras brasileiras e suas consequências no atual cenário sociolinguístico do português brasileiro. Na tentativa de compreender tal fato, voltaremos ao texto de Lobo (2015), com que iniciamos esta seção, visto que ela busca respostas verticalizadas em torno de 03 (três) políticas linguísticas articuladas na tentativa de extinguir a diversidade linguística. Vejamos a *Figura 1*:

**Figura 1** – Políticas linguísticas – tentativa de extinção da diversidade linguística do território brasileiro



Fonte: Elaborada pela autora, a partir da proposta de Lobo (2015).

Esses 03 (três) contextos sócio-históricos, correlacionados a políticas linguísticas, espelha a atual situação do Brasil,

[...] um país cuja população é hoje esmagadoramente monolíngue em português — ou em português brasileiro —, mas tal condição, além de não ter sido obra do acaso, teria, segundo Rosa Virgínia, sido definida a partir da segunda metade do século XVIII, quando «o multilinguismo 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 2004c: 20) (Lobo, 2015, p. 71).

Então, podemos inferir que a primeira e glotocida política linguística inicia-se na história em torno dos primeiros anos de colonização, quando os jesuítas se dedicam na aquisição dos “gentios” para mais tarde terem condições de convertê-los à fé cristã, e tem seu apogeu com

o Diretório dos Índios ou Diretório do Marquês de Pombal, de meados do século XVIII, o qual pôs fim ao trabalho missionário dos religiosos jesuítas nos aldeamentos indígenas, elevando-os à condição de vilas a ser administradas por um Diretor, cuja principal obrigação seria a implementação de um projeto dito civilizatório, o de civilizar os índios, devendo, para tal e primeiramente, extirpar-lhes as línguas de origem, fossem elas quaisquer línguas indígenas ou, em particular, a chamada geral (Lobo, 2015, p. 71).

Em 1757, com o Marquês de Pombal, se define explicitamente para o Brasil uma política linguística e cultural que fez mudar de rumo a trajetória que poderia ter levado o Brasil a ser uma nação de língua majoritária indígena, já que os dados históricos informam que uma língua geral de base indígena ultrapassara de muito as reduções jesuíticas e se estabelecia como língua familiar no Brasil eminentemente rural de então. O Marquês define o português como língua da colônia, conseqüentemente obriga o seu uso na documentação oficial e implementa o ensino leigo no Brasil, antes restrito à Companhia de Jesus, que foi expulsa do Brasil (Mattos e Silva, 1993, p. 83)

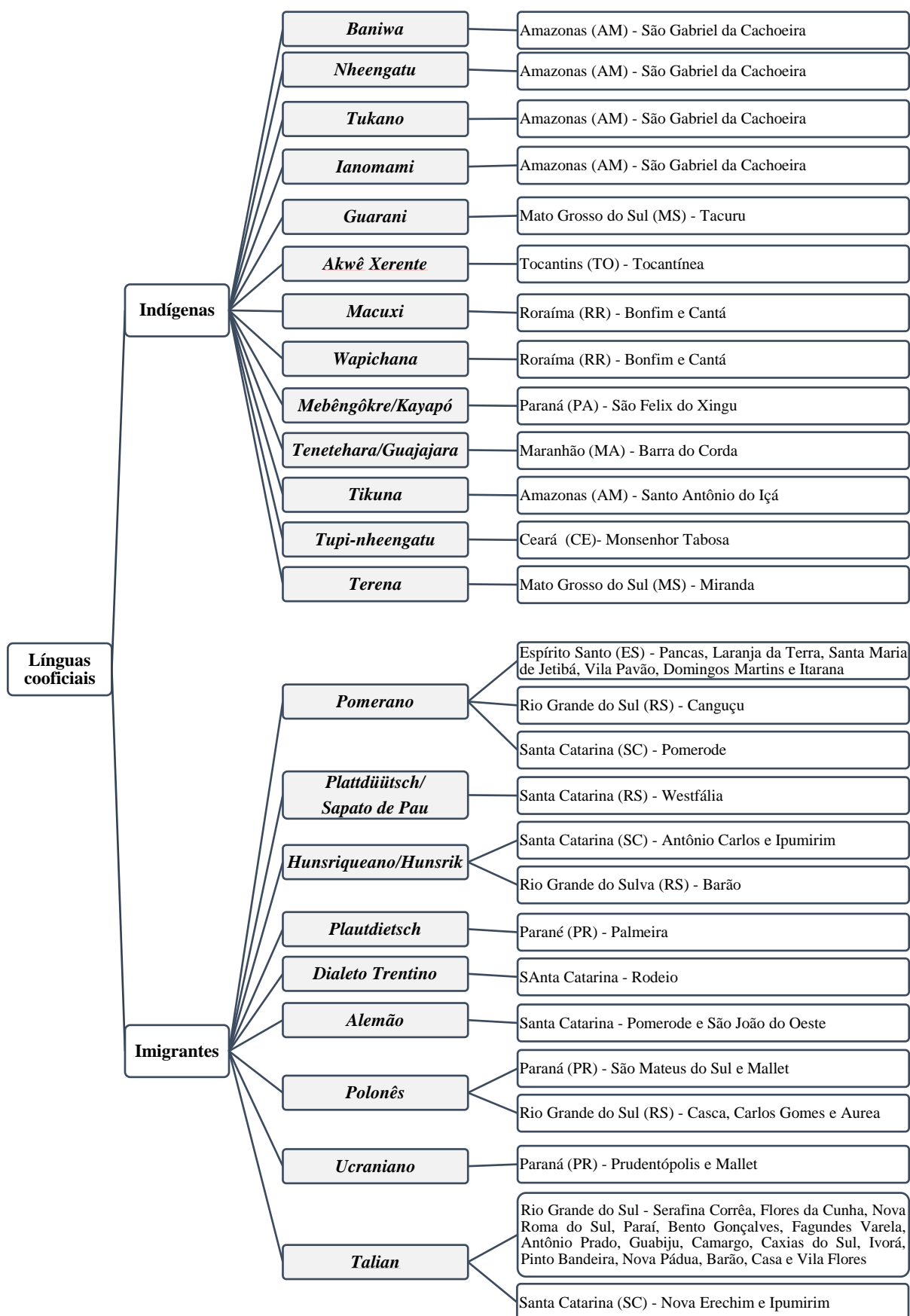
A segunda política linguística intitulada por mesclagem etnolinguística configura-se por uma perversa separação entre os coétnicos e colíngues africanos. Esse contexto

tem sido destacada por antropólogos, historiadores da escravidão e da história linguística do Brasil como um dos fatores para, das 200 a 300 línguas transplantadas com o tráfico de escravos, nenhuma ter-se implantado e difundido em território brasileiro, embora, a este respeito, se deva referir a ressalva de que algumas delas parecem ter tido uma relativa sobrevida, como teria sido o caso, por exemplo, do quimbundo, na Bahia do século XVII, e do ioruba, também na Bahia, mas já no século XIX (cf. Petter 2006: 124) (Lobo, 2015, p. 72).

Por fim, a terceira política linguística, apontada por Lobo (2015), direciona-se às línguas da imigração, também alvo do processo de imposição do português em terras brasileiras, visto que, sobretudo a partir de meados do século XIX, são “[...] atingidas não só, mas significativamente, pela política de nacionalização do ensino implementada pelo Estado Novo (1937-1945), que proibia o ensino escolarizado em qualquer outra língua que não fosse o português. Sucessivos decretos presidenciais [...] deram cabo das escolas de comunidades de alemães, italianos, poloneses e japoneses no Brasil” (Lobo, 2015, p. 72).

Diante deste percurso *etnocida e glotocida* (Mattos e Silva, 2004), Lobo (2015) retrata ainda algumas línguas que sobrevivem/resistem em terras brasileiras e conseguiram recentemente ser reconhecidas como cooficiais. Vejamos o *Figura 2*:

**Figura 2** – Línguas cooficiais em municípios brasileiros



Fonte: Elaborada pela autora, a partir dos dados do IPOL (2023).

É inegável que “[...] o processo do encontro entre a língua portuguesa, língua de dominação, com muitas línguas autóctones e as diversas línguas aqui chegadas, primeiro as africanas, depois as línguas de imigrantes, que tornaram esta área americana, multilíngüe de origem, ainda mais complexa lingüisticamente” (Mattos e Silva, 2004, p. 11). Daí a importância de descrevermos, ainda que sinteticamente, o passado sócio-histórico-lingüístico do Brasil, visto que, através dele, torna-se possível compreendemos o português brasileiro – não apenas heterogêneo e variável, mas também plural e polarizado (Lucchesi, 1994).

Na tentativa de finalizarmos esta seção, não as discussões suscitadas por esta temática, apresentaremos sinteticamente as proposições postas no âmbito da linguística para uma possível periodização da história da língua portuguesa no Brasil. No entanto iniciamos frisando que

As propostas de periodização da história de uma língua podem basear-se em fatores internos – mudanças que a estrutura da língua sofre – ou, mais frequentemente, em fatores externos – mudanças sócio-históricas supostamente correlacionadas às mudanças estruturais. Em ambos os casos, os critérios priorizados pelos historiadores das línguas são amplamente variáveis (Lobo, 2003, p. 396).

No *Quadro 2*, é possível observar uma descrição de algumas propostas de periodização para a língua falada no Brasil – Silva Neto, Marlos Pessoa de Barros, Tania Lobo, Volker Noll e Dante Lucchesi – que suscitam “[...] uma reflexão mais geral sobre a história da língua no Brasil, buscando compreender não apenas a formação da realidade linguística, mas também os fatores que lhe foram determinantes [...]. Quanto mais explícita for essa relação, mais a periodização alcançará seu objetivo maior: balizar a compreensão de conjunto da história que divide e orientar as pesquisas empíricas no campo” (Lucchesi, 2017, p. 348-349).

**Quadro 2** – Proposta para periodização da história linguística do Brasil

		PRIMEIRA FASE	SEGUNDA FASE	TERCEIRA FASE	QUARTA FASE	QUINTA FASE
SERAFIM DA SILVA NETO (1963[1951])	<b>Datação</b>	De <b>1532</b> , início da colonização portuguesa, até <b>1654</b> , data da expulsão dos holandeses do Brasil.  “é escassíssimo o elemento branco, que se vê afogado na maré de índios e africanos” (p. 74)	De <b>1654</b> até <b>1808</b> , data da chegada do Príncipe Regente da corte portuguesa ao Brasil.  “o elemento indígena, incompatível com a civilização europeia, vai rareando e desaparecendo, ao mesmo passo que cresce a influência de brancos e negros” (p. 80)	De <b>1808</b> até a atualidade (no caso, <b>1951</b> , data da publicação de sua proposta). Fase marcada pela chegada da Corte Portuguesa ao Rio de Janeiro.  “a dualidade linguística entre a nata social, viveiro de brancos e mestiços que ascenderam, e a plebe, descendentes dos índios, negros e mestiços da colônia” (p. 88-89)	-	-
	<b>Descrição</b>	A primeira proposta apresentada sobre a periodização linguística do português, em que o autor parte essencialmente de fatos da história política do Brasil – o início da colonização, a expulsão dos holandeses e a chegada da família real portuguesa.				
	<b>Crítica</b>	A maior crítica refere-se à fixação do final da primeira fase, visto que, como afirma Lobo (2003, p. 400), retomando a crítica de Marlos Pessoa (2003), “[...] o destaque para a ocupação holandesa não se justifica em uma proposta de periodização da história linguística brasileira, na medida em que a presença holandesa no Brasil, tendo durado apenas vinte e quatro anos, não teve consequências no plano da linguagem, ou seja, não provocou qualquer ‘ruptura do desenvolvimento linguístico normal que se dava na perspectiva das línguas indígenas, africanas e do próprio português.’”  Volker Noll emite o mesmo juízo (2008, p. 263), além de criticar que “as fronteiras temporais alegadas” são “todas externas”.				



MARLOS PESSOA (2003)	<b>Datação</b>	De <b>1534</b> , da divisão do país em capitânicas hereditárias, até <b>1750</b> , com a descoberta do ouro em Minas Gerais e a modernização do Estado português, com as reformas pombalinas. Se caracterizaria pela formação das variedades regionais e pelo multilinguismo.	De <b>1750 a 1922</b> , dividida em três subfases, que correspondem às respectivas etapas da formação do português comum brasileiro: <ul style="list-style-type: none"> <li>o de 1750 a 1808 - transferência da família real para o Brasil – “fermentação” do PB;</li> <li>o de 1808 a 1850 - fim do tráfico negreiro - “efetiva formação” do PB;</li> <li>o de 1850 a 1922 - eclosão do Movimento Modernista - “estabilização” do PB.</li> </ul>	A partir de <b>1922</b> , o fim do predomínio das oligarquias e o surto industrial. Se caracteriza pelo “estágio de elaboração da língua literária”.	-	-
	<b>Descrição</b>	Esta proposta procura articular fundamentos socioeconômicos com processos linguísticos de formação de uma língua comum, sobrepondo-se ao regionalismo inicial, e de elaboração de uma língua literária, que estariam relacionados aos fenômenos sociais da constituição do Estado Nacional e da urbanização.				
	<b>Crítica</b>	Segundo Lucchesi (2017, p. 352-353), fixar o fim da primeira fase em 1750, com base na descoberta das minas, é problemático, já que isso ocorreu no final do século XVII, ou seja, mais de cinquenta anos antes. Além disso, para se contrapor ao marco de 1922 proposto por Pessoa, Lobo defende a necessidade de considerar as transformações estruturais, invocando a proposição de Tarallo (1993) de que mudanças linguísticas teriam feito emergir uma gramática brasileira distinta da portuguesa, na passagem do século XIX para o século XX.				
TÂNIA LOBO (2003)	<b>Datação</b>	<b>até 1850:</b> multilinguismo generalizado; não urbanização; não escolarização e não estandardização linguística.	<b>após 1850:</b> multilinguismo localizado; urbanização; escolarização e estandardização linguística	-	-	-
	<b>Descrição</b>	Como afirma Lobo (2003), sua proposta é um tanto econômica e está fundamentada em 03 (três) aspectos, a saber: a história demográfico-linguística brasileira; o crescimento populacional associado ao processo de urbanização do país; e o processo de escolarização associado ao processo de estandardização linguística.				

	<b>Crítica</b>	Segundo Lucchesi (2017, p. 353), a contribuição dessa proposta é considerar primordialmente processos sociolinguísticos, socioeconômicos e culturais: a passagem do multilinguismo ao unilinguismo majoritário, a transformação de um país rural em um país urbano, a alfabetização e a normatização da língua. Contudo, sua divisão dicotômica aponta para dois estados estanques, não capturando o processo de transição essencial a todo processo histórico. Outros problemas decorrem da escolha da data de 1850 como marco divisório.				
<b>VOLKER NOLL (2008)</b>	<b>Datação</b>	De <b>1500</b> a <b>1550</b> , fase <b>inicial</b> : traslado da língua portuguesa para o Brasil	De <b>1550</b> a <b>1700</b> , primeira fase <b>formativa</b> : formação das primeiras características da língua portuguesa no Brasil	De <b>1700</b> a <b>1800</b> , fase <b>diferenciadora</b> : formação do português brasileiro, diferenciação das variedades europeia e brasileira.	De <b>1800</b> a <b>1950</b> , fase de <b>desenvolvimento da escrita e do ensino</b> : introdução da impressão tipográfica e da imprensa, implantação do ensino público oficial e criação do ensino superior; diferenciação progressiva da norma europeia.	De <b>1950</b> ao <b>presente</b> , fase de <b>nivelação</b> : evolução dos meios de comunicação, introdução da televisão, urbanização progressiva.
	<b>Descrição</b>	Proposta sugere uma periodização pautando-se em fatos históricos e, sobretudo, em critérios linguísticos – diferentemente das anteriores que se limitam quase exclusivamente a fatos históricos.				
	<b>Crítica</b>	Segundo Lucchesi (2017, p. 355-356), a proposição da fase inicial de 1500 a 1550 não faz qualquer sentido como transplantação da língua, já que a presença portuguesa no Brasil se resumia, no período, a alguns poucos indivíduos. Além disso, Noll ignora o multilinguismo que caracteriza o Brasil na época, destoando de todas as demais propostas, que tomam esse aspecto para definir essa fase inicial. Já a fase “diferenciadora” (1700-1800) coincide também com um grande avanço da língua portuguesa no território brasileiro (LUCCHESI, 2006), mas isso é ignorado por Noll, que se restringe à diferenciação entre o PE e o PB. Contudo, na definição das últimas fases, deve-se registrar o aspecto positivo de considerar processos sociais e culturais que afetam diretamente os usos linguísticos. Sendo assim, seu maior problema é basear-se em dois princípios impróprios, que estão interligados: (i) a língua evolui em função de sua estrutura interna, seguindo uma deriva; e (ii) o português não foi afetado pelo contato entre línguas no Brasil. O grande equívoco de Noll (2008) em relação ao papel do contato entre línguas na história sociolinguística do Brasil é se prender à dicotomia crioulização ou deriva. Para além da fragilidade de seus fundamentos teóricos, a proposta de Noll (2008) exhibe também um problema epistemológico, quando afirma que “como se trata de processos linguísticos que costumam se estender por décadas, não é razoável delimitar as fases por datas fixas, como se faz para os eventos históricos”. Porém Noll (2008, p. 269-277) prefere “uma classificação em séculos e semi-séculos” e propõe fases “na evolução do português brasileiro”:				

<b>LUCCHESI (2017)</b>	<b>Datação</b>	<p><b>Tupinização da Conta:</b> se inicia <b>1000</b>, com a expansão tupi pelo litoral, e se encerra no ano de <b>1532</b>, com o efetivo início da colonização do Brasil pelos portugueses;</p>	<p><b>Multilinguismo generalizado:</b> de <b>1532 a 1695</b>, é marcado pelo plurilinguismo, no qual o português se insere, de forma minoritária, em um extenso mosaico de centenas de línguas indígenas faladas no interior do país.</p>	<p><b>Homogeneização Linguística:</b> inicia-se em <b>1695</b>, com a descoberta do ouro em Minas Gerais, e se estende até a Revolução de <b>1930:</b></p> <ul style="list-style-type: none"> <li>○ 1695-1808 – Primeira vaga de aportuguesamento;</li> <li>○ 1808-1850 - Segunda vaga de aportuguesamento;</li> <li>○ 1850-1930 - Terceira vaga de aportuguesamento</li> </ul>	<p><b>Nivelamento linguístico:</b> de <b>1930 aos dias atuais</b>, é marcada pela crescente industrialização e urbanização do país, pela difusão da norma urbana culta, apagando as marcas do contato na norma popular.</p>	
	<b>Descrição das fases</b>	<p>Povos tupis expulsam povos primitivos, tornando o tupi e o tupinambá as línguas mais faladas no litoral brasileiro.</p>	<p>Baixa densidade demográfica da sociedade colonial faz com que o português conviva com as línguas gerais indígenas e línguas francas africanas. No interior do país, ocorrem eventualmente processos de pidginização/crioulização do português, e centenas de línguas indígenas se conservam nas tribos não contatadas.</p>	<p>Imposição do português como língua hegemônica, restringindo fortemente o multilinguismo.</p>	<p>Industrialização e urbanização promovem a difusão da norma urbana culta, apagando as marcas do contato na norma popular; a norma culta se afasta do padrão lusitanizado; porém concentração de renda mantém a polarização sociolinguística e fomenta o preconceito linguístico.</p>	
	<b>Descrição da Proposta</b>	<p>Segundo Lucchesi (2017, p. 376), esta proposta busca integrar os processos linguísticos na matriz mais ampla dos processos sócio-históricos. Deve-se destacar que ela se fundamenta exclusivamente em processos sociolinguísticos: a tupinização da costa, o multilinguismo generalizado, a homogeneização linguística e o nivelamento linguístico. Ao traçar esse liame de processos sociolinguísticos ao longo de todas as fases, esta periodização favorece a compreensão global da formação histórica da realidade sociolinguística do Brasil, podendo impulsionar a pesquisa empírica no campo.</p>				

Fonte: Elaborada pela autora, a partir de Lucchesi (2017).

Nesta tese, assumimos a perspectiva proposta por Lucchesi (2017), que propõe uma perspectiva inovadora, a partir da visão da história sociolinguística do Brasil, tomando como base a Teoria da Variação Linguística, que compreende que os processos de mudança linguística não podem ser compreendidos fora do contexto social em que ocorrem. Ainda frisamos a importância de termos traçado uma descrição em relação ao processo de colonização em terras brasileiras, na tentativa de demonstrar a passagem de “um multilinguismo generalizado a um multilinguismo localizado” (Mattos e Silva, 2004, p. 14), apresentando, a partir de Lobo (2015), as políticas linguísticas brasileiras que colaboraram para a assertiva proposta por Mattos e Silva (2004). Por fim, encerramos esta seção apresentando as propostas de periodização, assumindo a de Lucchesi (2017), por motivos já mencionados e por acreditarmos, juntamente com o autor, que “[...] uma abordagem que não considera o contato é defectiva, em face das estreitas relações que uniam colonizadores e escravos africanos, seja no universo urbano, seja no meio rural” (Lucchesi, 2017, p. 350). Assim, seguimos, na seção posterior, uma breve descrição das línguas africanas acolhidas forçosamente em terras brasileiras, visto sua importância no processo de formação do português brasileiro, bem como, dada a proposta assumida nesta tese, um estudo do presente para o passado, através da constituição e análise de *corpus* oral: o que amostras orais do português popular de comunidades rurais afro-brasileiras podem nos dizer sobre o presente e, de certa forma, sobre o passado do português popular do Brasil?.

## 2.2 Línguas africanas no Brasil

Dada a importância da relação entre as línguas africanas e o português brasileiro, especificamente entre os estudos linguísticos contemporâneos desenvolvidos no Brasil<sup>6</sup>, nesta seção, apresentaremos as línguas africanas chegadas ao Brasil, na tentativa de enfatizar que os africanos e afrodescendentes foram os principais difusores da língua portuguesa no Brasil e os principais formatadores do PB em sua variante social majoritária — o português popular brasileiro — proposta já descrita na seção anterior, a partir de Lobo (2015).

Sendo assim, cientes de que a formação da nossa língua está inteiramente interligada ao processo de colonização brasileiro, quando ocorre o encontro plurilinguístico entre o

---

<sup>6</sup> Borges (2015) apresenta uma inédita bibliografia de referência, com o levantamento de 322 textos publicados no Brasil, entre os anos de 1870 e 2010, sobre a relação do PB com as línguas africanas ao longo dos séculos XIX, XX e XXI. Disponível em: [https://www.teses.usp.br/teses/disponiveis/8/8139/tde-06082015-114127/publico/2015\\_PatriciaDeSouzaBorges\\_VCorr.pdf](https://www.teses.usp.br/teses/disponiveis/8/8139/tde-06082015-114127/publico/2015_PatriciaDeSouzaBorges_VCorr.pdf)

português europeu e as muitas línguas indígenas e africanas, trataremos, neste momento, ainda que brevemente, das línguas africanas que possivelmente marcam a formação do PB<sup>7</sup>, visto que

um dos pontos nevrálgicos do debate tem se estabelecido em torno do aporte linguístico africano introduzido no Brasil, tendo em vista que, entre o final do século XVI e o início do século XIX, os africanos representavam entre 20% e 30% de todo o contingente populacional em território brasileiro, chegando a superar numericamente outros grupos (índios, portugueses, brasileiros descendentes de africanos e europeus, entre outros) ao longo do século XVII – ver Mussa (1991:163) (Avelar; Galves, 2014, p. 243).

Na tentativa de compreender esse cenário posto ao PB, Alencastro (2009) afirma que, para entendermos a presença das línguas africanas no atual idioma brasileiro, é preciso situar os povos africanos na formação do Brasil, como fizemos na seção anterior, e situar o continente africano. Afinal, a África plural, dotada de uma grande diversidade de povos, culturas e línguas, com suas correspondentes variedades dialetais, configura-se como um dos caminhos que nos faz entender mais sobre o português brasileiro.

Nesse sentido, iniciamos afirmando que “o reconhecimento das línguas africanas decorre da descoberta progressiva da África, no século XV, quando a Europa se volta para ela, movida por razões mercantilistas, envolvidas num apostolado missionário” (Petter, 2017, p. 28), em um contexto de exploração sistemática do continente, através dos processos de escravidão, tráfico e colonização.

A História nos desvela que, no período das conquistas marítimas lusitanas, no século XV, muitas áreas da costa africana foram conquistadas, enquanto Portugal buscava estabelecer seu comércio de especiarias com a Índia. Segundo Mingas (2000), o processo de colonização da África foi resultado de um jogo de interesses europeu, o que gerou a partilha do continente, na Conferência de Berlim<sup>8</sup>, em 19 de novembro de 1884 e 26 de fevereiro de 1885, momento em que hipoteticamente buscaram organizar a ocupação do continente.

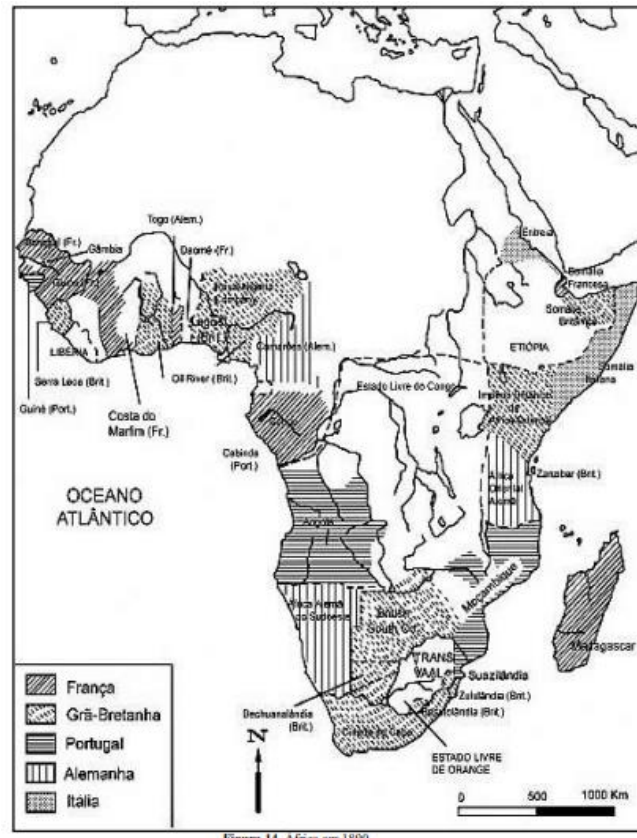
Os estudos sobre a colonização do continente africano afirmam que a Conferência resultou na partilha da África, uma divisão arbitrária que desrespeitou as características étnicas e culturais de cada tribo que vivia naquele território, muitas tribos rivais foram obrigadas a conviver em um mesmo espaço geográfico, e outras unidas foram obrigadas a se separar, culminando em

<sup>7</sup> Sugerimos a leitura do artigo “O papel das línguas africanas na emergência da gramática do Português Brasileiro” (Avelar; Galves, 2014), em que são postos argumentos favoráveis à hipótese de que as línguas africanas faladas pelos escravos introduzidos no Brasil desempenharam um papel significativo na emergência da gramática do português brasileiro. Os autores apresentam um paralelismo morfossintático entre o português brasileiro e o português africano, bem como entre essas variedades e as línguas bantas, para propor que a aquisição do português como segunda língua pelos africanos produziu mudanças em duas direções: (i) transferência de propriedades sintáticas das suas línguas maternas para o português em formação no Brasil e (ii) reestruturação desencadeada pela dificuldade no aprendizado de marcas gramaticais específicas do português.

<sup>8</sup> A Conferência de Berlim foi um fórum de representantes de diversos países que se reuniram na cidade de Berlim, capital do Império Alemão, entre 15 de novembro de 1884 e 16 de fevereiro de 1885 para determinar questões relacionadas à ocupação dos territórios africanos pelas potências imperialistas europeias (BARBOSA, 2008).

muitos conflitos, sendo a história desses povos anulada diante do etnocentrismo europeu, que cria condições para divisão da África entre as potências europeias (cf. Figura 3); e na criação dos Estados africanos no seu atual formato. Sem dúvidas, como afirma Barbosa (2008, p. 05), uma das consequências mais nefastas desses acordos foi o estabelecimento das condições políticas e ideológicas das “guerras étnicas” africanas<sup>9</sup>.

**Figura 3 – África em 1890**



Fonte: Barbosa (2008, p. 196).

Importante frisarmos este momento no tempo, pois, como afirma Anjos (2011)<sup>10</sup>, uma das questões estruturais direcionadas à cultura africana no Brasil e, acrescentamos, à história

<sup>9</sup> Indicamos a leitura da tese de Frederico Antonio Ferreira (2018), intitulada *No festim dos lobos: o império do Brasil e o colonialismo na África centro-ocidental (1880-1885)*, em que, numa perspectiva, apresenta o posicionamento da política externa brasileira quanto à Conferência de Berlim (1884-1885) e o avanço neocolonialista dos países europeus sobre o litoral da África Centro-Occidental. Portanto, através dos documentos do Arquivo Histórico do Itamaraty e de diversos jornais e revistas circulantes no Rio de Janeiro na primeira metade da década de 1880, o autor busca elucidar as continuidades e rupturas nas relações sociais, culturais, econômicas e políticas entre o Império do Brasil e as diversas comunidades políticas do litoral africano neste período. Disponível em: <https://tede.ufrj.br/jspui/bitstream/jspui/4747/2/2018%20-%20Frederico%20Antonio%20Ferreira.pdf>

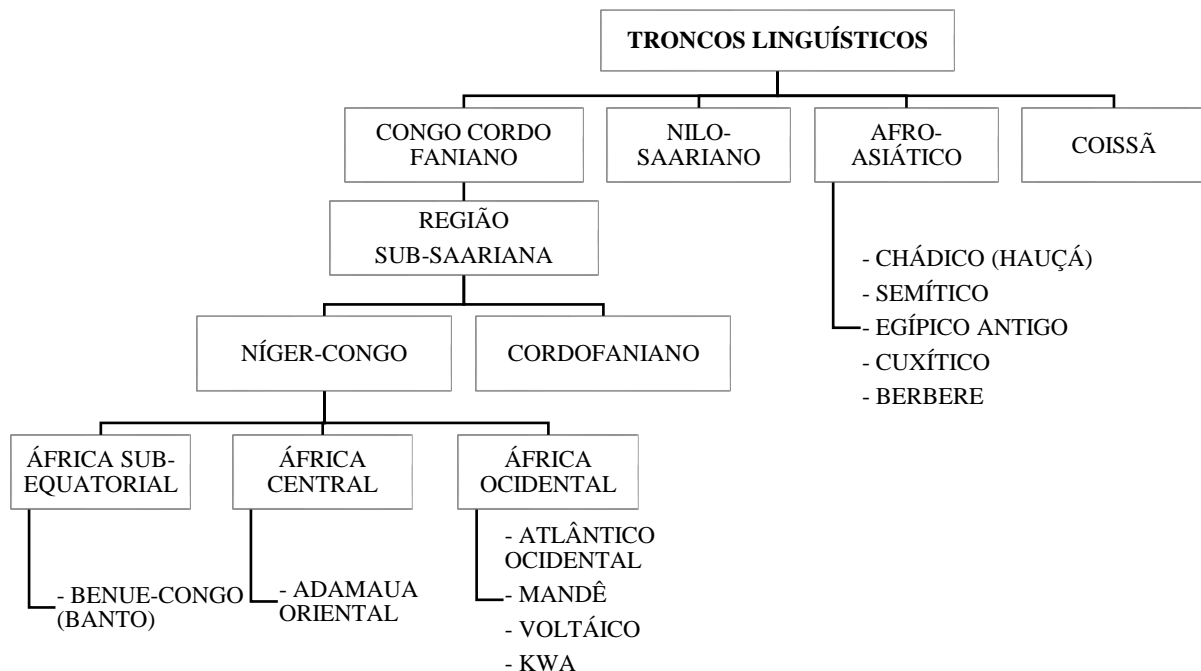
<sup>10</sup> Sugerimos a leitura do artigo *Cartografia da diáspora África – Brasil* (2011), em que Anjos resgata as principais referências da diáspora África-América-Brasil e configura a expressão espacial da população afro-brasileira contemporânea. Disponível em: <https://ojs.ufgd.edu.br/index.php/anpege/article/view/6570>

linguística africana, está relacionada aos aspectos historiográficos e geográficos da formação e distribuição da população de ancestralidade na África do território brasileiro. Nesse sentido, não podemos deixar de relatar que o tráfico de negros africanos foi um dos maiores movimentos populacionais da história e certamente a maior migração forçada por mar antes da grande emigração europeia, também originariamente para as Américas, que se desenvolveu à medida que o tráfico de escravos no Atlântico conhecia o seu fim (Fage, 2010, p. 269).

O maior processo de migração forçada da história da humanidade representa as muitas Áfricas espalhadas, em especial pelas Américas; dentre elas, as que se fizeram e se fazem presentes no Brasil através dos seus povos africanos e afrodescendentes. Dentre as Áfricas, aqui tentaremos, a partir de agora, centrar o olhar no plurilinguístico africano chegado ao Brasil através desse infame comércio transatlântico, que faz nascer no Brasil uma língua plural, que marca em sua estrutura linguística, não apenas lexical, a identidade de seus povos – portugueses, africanos, indígenas e imigrantes.

Ao se tratar das línguas africanas, estudiosos as distribuíram em relação às proximidades ou afastamentos, organizando-se em famílias de línguas. Castro (2005), a partir da classificação de Greenberg (1963), apresenta a seguinte distribuição em troncos linguísticos:

**Figura 4** – Distribuição de Greenberg: troncos linguísticos africanos



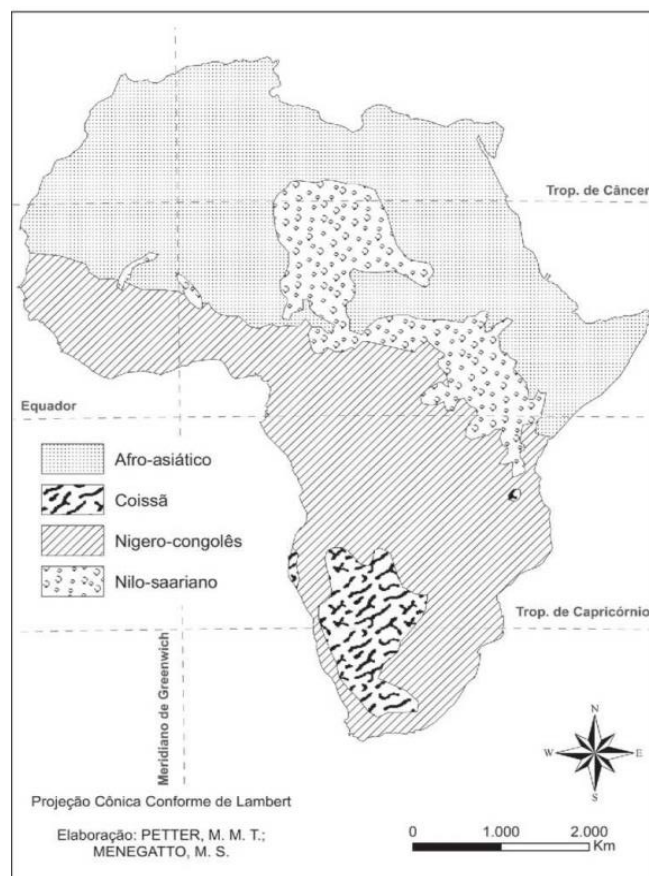
Fonte: Castro (2005, p. 27)

O esquema nos permite visualizar que Greenberg (1963) subdivide as línguas africanas em quatro troncos linguísticos, sendo um deles o *Congo Cordofaniano*, que representa a região

subsaariana e se subdivide em duas grandes famílias, a *Níger-Congo* e a *Cordofaniano*: a primeira se faz presente em três regiões – *África Subequatorial, Central e Ocidental* –, sendo o grupo banto representado pela grande família *Níger-Congo*, no ramo *Benue-Congo*. A repartição, proposta por Greenberg (1963), foi adotada por grande parte dos linguistas africanistas, com algumas revisões ligadas à reorganização interna de cada tronco (Bonvini, 2009, p. 23).

A proposta apresentada na *Figura 4* foi a primeira a considerar todo o continente africano e, até então, a que melhor descreveu a metodologia para o estudo das línguas africanas. No entanto atualmente temos outras propostas, a citar a de Heine e Nurse (2000), que reconhecem 04 (quatro) troncos – Nígero-congolês, Nílo-saariano, Coissã e Afro-asiático – que coincidem, em grande parte, com os indicados por Greenberg (1963) (*cf.* *Figura 4*), e diferindo especificamente em relação à divisão interna dos grupos (Petter, 2015). Vejamos a distribuição geográfica desses 04 (quatro) troncos linguísticos:

**Figura 5** – Línguas africanas: troncos linguísticos



Fonte: Petter (2015, p. 52), adaptado de Heine e Nurse (2000).



Em seu capítulo *A classificação das línguas da África* (2015), Petter traça sistematicamente a história da classificação dos troncos linguísticos e os pontos divergentes entre as 02 (duas) propostas e afirmam que “muito embora os estudos sobre a classificação genética de línguas africanas estejam avançados, ainda há algumas controvérsias sobre as grandes divisões em troncos” (Petter, 2015, p, 84). No entanto as propostas caracterizam-se como uma tentativa de revelar o parentesco entre as línguas de seus povos, bem como a relação mútua entre os diversos aspectos culturais, afinal “a língua é um fenômeno histórico e um extraordinário veículo de informações de todo o gênero, não só de ordem estritamente linguística, mas também de cunho extralinguístico, remetendo ao contexto histórico mais amplo” (Petter, 2015, p. 50).

Ainda em relação à classificação das línguas africanas, Petter e Araújo (2015, p. 45) apontam as teorias linguísticas que se debruçam sobre línguas presentes no continente africano, a saber:

- (i) Teorias sintáticas e fonológicas;
- (ii) Teorias da gramaticalização e línguas em contato;
- (iii) Linguística Areal e tipologia linguística.

Dentre essas destacamos a (iii), visto que considera as cerca de 2.000 línguas do continente africano e se interessa “[...] numa classificação das línguas a partir das formas com que cada língua codifica formalmente uma determinada categoria linguística” (Petter; Araújo, 2015, p. 44). Vejamos a proposta de traços tipológicos apresentada pelas autoras:

**Quadro 3** – Traços tipológicos das línguas africanas

TRAÇOS RESTRITOS À ÁFRICA	TRAÇOS MAIS COMUNS À ÁFRICA
Cliques	Oclusivas lábio-velares
Flaps labiais	Implosivas ou oclusivas não obstruintes
Vários tipos de harmonia vocálica	Harmonia vocálica baseada em ATR
Ideofones para distinção de cor	Oclusivas pré-nasalizadas no início de palavras
Distinção de caso expressa exclusivamente por tom, normalmente para o nominativo	Sistemas de classes nominais
Não obrigatoriedade de concordância de verbos transitivos com seus objetos	Sistema de casos nominativos marcados
	Marcação da negação ao final da sentença
	Pronomes logofóricos
	Marcação de foco por meio da flexão verbal
	SOVX como ordem de palavras básicas

Fonte: Organizado a partir de Petter e Araújo (2015, p. 45).

A apresentação do *Quadro 3* é um instigante convite à pesquisa no campo linguístico, a possíveis comparações dos traços tipológicos recorrentes nas línguas africanas e os que ainda se fazem sincronicamente presentes no português brasileiro, especialmente em sua vertente popular. Assim, Petter e Araújo (2015, p. 45) afirmam que há propriedades que são muito mais comuns na África do que em outros continentes e que “qualquer avaliação de traços tipológicos em determinado grupo de línguas ou nas línguas do mundo se mostrará incompleta se for desconsiderado o continente africano, não apenas pelo número de línguas que ele engloba, mas também pela diversidade de traços tipológicos considerados típicos desse continente”.

Dadas essas pequenas considerações, direcionaremos nosso olhar às línguas africanas trazidas forçosamente pelo tráfico negreiro ao Brasil, cientes de que essa cruel comercialização de pessoas e falantes de muitas línguas teve início, no Brasil, em 1502, e foi oficialmente extinta em 1860, transplantando cerca de 3.500.000/3.600.000 africanos (Petter, 2015), distribuídos por estudiosos, a citar Mattoso (2003, p. 22-23) e Bonvini (2009, p. 26), que se retratam a proposta de Luiz Viana Filho, em 04 (quatro) grandes ciclos de importação de escravos:

**Quadro 4** – Ciclos do tráfico negreiro e origem dos escravos traficados

SÉCULO	CICLO	ORIGEM ÉTNICA	DESTINO
XVI	Ciclo da Guiné	Negros da Costa da Guiné.	Pernambuco, Bahia e Rio de Janeiro (demanda da economia açucareira)
XVII	Ciclo de Angola e do Congo	Negros banto, oriundos da África equatorial e central	Pernambuco, Bahia e Rio de Janeiro (demanda da economia açucareira e conflitos luso-holandeses)
XVIII	Ciclo da costa de Mina e, posteriormente, da Baía de Benin	Negros Sudaneses	Minas Gerais (extração aurífera)
XIX	Tráfico ilegal de escravos	Negros banto (Angola e Moçambique)	São Paulo (economia cafeeira)

Fonte: Mattos e Silva (2003, p. 22-23).

O panorama secular apresentado no *Quadro 4* demonstra que os africanos escravizados trazidos para o Brasil apresentavam uma composição étnica relativamente coesa, com relação aos ciclos estabelecidos, fato que se reflete na sua caracterização linguística. Nesse sentido, Bonvini (2009, p. 27) afirma que a repartição em ciclo só deve ser entendida dentro das suas grandes linhas, porque o fim de um ciclo não implica o interrompimento do tráfico de negros vindo da região representada no ciclo anterior. Nesse cenário – brasileiro –, Angola se caracteriza como o grande centro de tráfico negreiro, e torna-se incontestável que a escravidão implicou a transferência de milhões de africanos, os quais endossaram o cenário sócio-histórico,

compondo um contexto de imposição e de plurilinguismo (Castro, 2005; Lucchesi, 2003, 2009; Mattos e Silva, 2004a; Mussa, 1991; dentre outros).

A situação de plurilinguismo, instaurada na costa da África, fica mais evidente quando se trata de Angola, uma vez que os dados de Teyssier (1997) comprovam que, dentre as repúblicas, Angola era a maior em dimensão territorial (1.246.700 km<sup>2</sup>) e a segunda maior em números de habitantes (5.840.000). Ao tratar de língua, não poderemos deixar de refletir, através das palavras de Mingas (2000), sobre o processo de ensino em Angola. Com a política portuguesa de ensino, tem-se a supremacia da língua portuguesa em detrimento das muitas línguas locais e a “aplicação prática deste tipo de política linguística foi apoiada por uma vigilância cada vez mais acentuada da polícia política portuguesa, a famosa PIDE, que via na utilização de qualquer uma das línguas locais ou mesmo de vestuários locais pelos ‘assimilados’, atitudes subversivas” (Mingas, 2000, p. 48). Portanto, a língua portuguesa era a única língua ensinada na escola primária, exceto as escolas dirigidas por missionários, em que a língua local podia ser usada como meio de ajuda para facilitar a aprendizagem do português. Com a política linguística, a língua portuguesa era uma espécie de passaporte à ascensão social, e aos “assimilados” era negado o direito de falar outra língua que não fosse português.

Albuquerque e Fraga Filho (2006, p. 45) comentam que o

[...] tráfico transatlântico promoveu o povoamento do Brasil por gente vinda de diversas regiões do continente africano. A metrópole portuguesa adotou a política de misturar escravos de diferentes regiões e etnias para impedir a concentração de negros da mesma origem na colônia, os quais, solidários na cultura e falando a mesma língua, podiam se rebelar mais facilmente. Essa política nem sempre era seguida à risca, pois a depender das relações comerciais na África, os traficantes tendiam a transportar escravos que em sua maioria vinham duma mesma região.

Foi por isso que, no século XVI, a maioria dos escravos trazidos para o Brasil vinha da região da Senegâmbia, denominada Guiné pelos portugueses[...]. Mas, no decorrer daquele século até a primeira metade do século XVIII, os chefes políticos e mercadores do território presentemente ocupado por Angola forneceram a maior parte dos escravos utilizados em todas as regiões do Brasil [...]. Luanda, Benguela e Cabinda eram os principais portos de embarque.

No que tange às línguas africanas vindas para o Brasil, Bonvini (2009) e Petter (1998) estimam entre 200 e 300 línguas, e Mendonça (2012) relata que, na linguística africana, só se projetam duas individualidades fortes: línguas bantu e línguas não bantu. Diante da afirmação de Mussa (1991, p. 146), segundo a qual “[...] o percentual de falantes bantu foi sempre superior, e quase sempre maciçamente, em todo o período do tráfico” no Brasil, pode-se deduzir que a influência bantu é muito mais profunda em razão da sua distribuição humana em território brasileiro (Castro, 2005, p. 6). Dessa forma, dentre os negros falantes de língua africana, habitantes do Brasil, temos:

**Tabela 3** – População de falantes das diversas línguas africanas que chegaram ao Brasil

	<b>XVI</b>	<b>XVII</b>	<b>XVIII</b>	<b>XIX</b>
<b>Oeste-atlântico</b>	20%	7%	1%	3%
<b>Mande</b>	20%	7%	1%	3%
<b>Kru</b>	14%	5%	1%	3%
<b>Gur</b>	1%	3%	8%	9%
<b>Benue-Kwa (não bantu)</b>	7%	10%	20%	24%
<b>Bantu</b>	35%	65%	64%	50%
<b>Outros</b>	3%	3%	5%	8%

Fonte: Mussa (1991, p. 163).

Aqui destacamos o grupo banto, considerado o maior e mais conhecido conjunto linguístico, atraindo os olhares de muitos estudiosos no século XIX (Petter, 2015). Castro (2005), ao tratar das línguas africanas, afirma que o termo banto foi proposto por W. Bleek, em 1862, na primeira gramática comparativa do banto, no intuito de denominar a família linguística que havia descoberto, “[...] composta de várias línguas oriundas de um tronco comum, o protobanto, falado há três ou quatro milênios atrás” (Castro, 2005, p. 25). A autora ainda afirma que só posteriormente pesquisadores se utilizaram desse termo para denominar os indivíduos que habitavam a extensão abaixo da linha do equador.

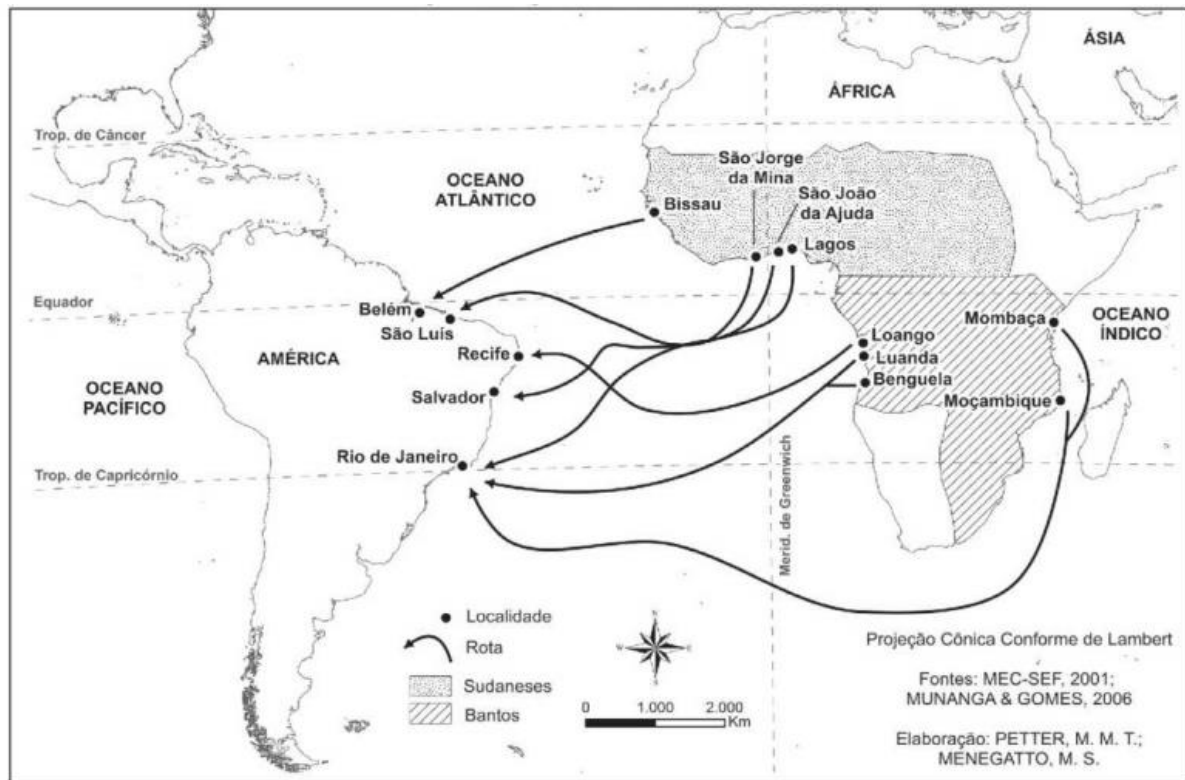
Estimam-se 400 mil, ou mais, línguas banto faladas em toda extensão territorial da África. Pesquisas comprovam suas semelhanças, tanto na estrutura gramatical quanto no vocabulário. No século XIX, com o avanço dos estudos científicos sobre as línguas banto, elas passam a se definir como uma família linguística, com muitas similaridades, dentre as muitas outras famílias linguísticas presentes na África. Dessa forma, “[...] qualquer língua banto podia ser parcialmente inteligível para os falantes de línguas próximas, e talvez não totalmente ininteligível para os falantes de qualquer outra língua banto, mesmo uma que fosse falada a uma distância considerável” (Fage, 2010, p. 31), o que leva alguns autores a considerar as línguas banto como uma família distante/separada das outras famílias linguísticas africanas.

Em um percurso cronológico sobre as origens da sociedade africana, Fage (2010, p. 31) afirma que “[...] os povos hoje chamados Banto se tornaram o tronco humano dominante em toda a metade meridional de África, à exceção do extremo sudoeste, dentro e à volta do deserto do Calaári [...]” e que o termo banto corresponde a uma classificação linguística das

línguas faladas, atualmente, pelos povos negros, em sua grande maioria, habitantes da metade meridional da África.

Para melhor compreensão das rotas linguísticas, vejamos o mapa que indica as regiões de onde partiram e as cidades a que foram levados os africanos durante o tráfico negreiro:

**Figura 5** – Principais rotas do tráfico transatlântico



Fonte: Petter (2015 *apud* Munanga, 2009).

No que tange aos dados demográficos, Alkmim (2007) relata que, ao longo do período colonial, o tráfico de escravos se centralizou nas regiões da África, porém pouco se sabe sobre o contingente de negros traficados para o Brasil, devido à ausência de documentos históricos, mas estima-se entre 3.500.000 e 3.600.000 escravos africanos, até 1850. Mattos e Silva (2004a, p. 105) ressalta que Robert Conrad – em sua pesquisa sobre os últimos anos de escravidão no Brasil – demonstra que, às vésperas da abolição da escravidão, em 1884, a população escrava se concentrava em Minas Gerais, Rio de Janeiro e São Paulo, graças à migração do Nordeste para a área cafeeira. Vejamos as estimativas:

**Tabela 4 – População escrava**

<b>PROVÍNCIA</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 Norte	60.000
Maranhão	45.000

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

A presença maciça de africanos e afrodescendentes na demografia histórica do Brasil demonstra não apenas a atuação dos escravizados frente à economia da colonização, a mobilidade geográfica, mas também o significado social e linguístico dos espaços ilegítimos de escravidão, os quais nos têm permitido inferir que esse segmento numeroso e operante – os africanos e afrodescendentes – é o agente principal da difusão do português no território brasileiro, na sua face majoritária, a popular ou vernácula (Mattos e Silva, 2004).

Esse cenário, apresentado no contexto brasileiro (*cf.* seção 2.1) e africano (*cf.* seção 2.2), serve para compreendermos, ainda que parcialmente, o processo de formação e difusão da língua portuguesa no Brasil, inferindo que as situações de aprendizagens impostas irregularmente por parte da elite colonizadora projetam-se hoje no português brasileiro, especialmente em sua vertente popular.

No entanto não podemos deixar de relatar que, como afirma Petter (2007), o conhecimento das línguas africanas foi constituído por um longo percurso histórico, legando documentos que não foram, ainda, exaustivamente estudados pela historiografia linguística, até porque só muito recentemente a Linguística Africana passa a ocupar cenários nas Universidades Brasileiras<sup>11</sup>, quando iniciamos os primeiros passos na tentativa de resgatarmos uma dívida histórica com nossos ancestrais africanos. Nesse sentido, é possível afirmar que

---

<sup>11</sup> Ainda são necessários avanços no que tange a “[...] centros ou institutos de pesquisas voltadas exclusivamente para a Linguística Africana, a exemplo dos países que possuem uma longa tradição nos estudos das línguas da África (Alemanha, França, Bélgica etc.). No Brasil, além da Universidade de São Paulo, a Fundação Universidade Federal de Rondônia (Unir) também agrega pesquisadores que tem se dedicado nos últimos anos a consolidar uma linha de pesquisa em Linguística Africana, nomeadamente sob um viés etnolinguístico e lexicográfico, principalmente relacionado ao domínio das línguas bantas” (Petter; Araújo, 2015, p. 46).

descrever as línguas do continente ainda é a meta principal da Linguística Africana. Nos últimos trinta anos, um esforço sem precedentes foi feito para descobrir *in loco* (linguística de campo) a singularidade dessas línguas através de materiais de primeira mão, numerosos e diversificados, testemunhas insubstituíveis de um estado de língua e da variedade de seus usos. Muitas línguas foram descritas e passou-se a identificar *Linguística Africana* com *Linguística Descritiva*. Se por um lado se valorizou um aspecto positivo com essa identificação – o trabalho descritivo –, por outro, ficou implícita, por algum tempo, uma crítica ao não comprometimento de alguns trabalhos como ‘teoria’. A Linguística Africana, hoje, não distingue descrição de teoria, porque considera impossível uma descrição sem teoria (Petter; Araújo, 2015, p. 28).

A história da Linguística Africana mostra-nos que a descrição das línguas africanas é praticamente indissociável da história da colonização da África e grande parte dos resultados das políticas linguística, hoje, nesse continente, está intimamente relacionada a esse passado colonial. Podemos observar que desde os primeiros autores e primeiras obras dedicadas ao conhecimento das línguas africanas até uma internacionalização dos estudos dessas mesmas línguas ainda há muito a ser feito no que concerne à descrição e à documentação das cerca de 2.000 línguas em todo o continente (Petter, 2015, p. 48).

Dada a importância da Linguística Africana, Petter e Araújo (2015) trazem a descrição da periodização da história da Linguística africana, a partir da proposta de Bonvini. Vejamos:

**Quadro 5** – Periodização da história linguística africana, por Petter (2017)

PERÍODOS		DATAÇÃO	DESCRIÇÃO
Período colonial	<b>Primeiro</b>	1440-1600	Descoberta do continente africano e o choque de diversidade das línguas da África; As línguas africanas como recursos práticos; A procura e a formação de intérpretes <sup>12</sup>
	<b>Segundo</b>	1600-1800	O continente africano encarado como potencial de mão de obra e de almas em perigo; As línguas africanas como instrumento de salvação e o latim como modelo de descrição.
	<b>Terceiro</b>	1800-1900	A descoberta e a conquista do interior do continente africano; A variedade linguística africana como obstáculo à ‘civilização’ colonizadora e o modelo indo-europeu como referência linguística
	<b>Quarto</b>	1900-2000	A diversidade das línguas e a universalidade da linguagem. Da colonização à descolonização/independência; A variedade das línguas africanas valorizada cientificamente, mas sob a ameaça de extinção: o modelo linguístico indo-europeu em revisão.

<sup>12</sup> “Os intérpretes eram escolhidos entre os escravos de alguns senhores africanos que os emprestavam aos portugueses para que eles aprendessem a língua portuguesa, em seguida eram trocados por um outro escravo que estava no navio; finalmente, eles seriam libertados quando um desses intérpretes conseguisse conquistar outros quatro escravos para os seus senhores. Esses intérpretes foram conhecidos como Negro trusimã [...] – ‘mestre de língua’ ou simplesmente ‘(o) língua’. Na impossibilidade de encontrar um ‘trusimã’, os negros eram levados à força para Lisboa, onde seriam confrontados com os numerosos intérpretes negros que estavam em Portugal e lhes seria ensinada a língua portuguesa.” (Petter; Araújo, 2015, p. 30)

<b>Período pós-colonial</b>	<b>Primeiro</b>	1960-1980	Desenvolvimento de trabalho no campo dos estudos linguísticos tanto na África quanto nos antigos países colonizadores; Desenvolvimento de uma metodologia de campo baseada em questionários de pesquisa linguística <sup>13</sup> ; Surgimento de escolas linguísticas <sup>14</sup>
	<b>Segundo</b>	1980-2000	Internacionalização da pesquisa africanista; Reavaliação das pesquisas anteriores; Renovação da abordagem das línguas.

Fonte: Organizado pela autora, a partir de Petter e Araújo (2015).

Não há dúvidas de que agora estamos no caminho que nos conduzirá ao preenchimento de lacunas seculares no que tange à história das línguas africanas em terras brasileiras, visto que a descrição e a documentação dessas línguas marcam a importância sócio-histórica do povo africano, bem como a compreensão em relação às situações de contatos linguísticos na formação do português falado no Brasil. Que propostas como a desenvolvida nesta tese, de constituição de *corpus* oral em comunidades afro-brasileiras, sejam um caminho colaborativo nesse contexto em que temos os africanos e os afrodescendentes como protagonistas de sua história e atuantes, não apenas lexicalmente, no processo de formação do português brasileiro.

### 2.3 Formação do Português Brasileiro: algumas hipóteses e discussões

A Linguística Africana no Brasil, citada na seção anterior, está relacionada às questões que envolvem o contato do português com as línguas africanas, visto que os primeiros contatos são muito antigos, quando as migrações para regiões distantes fizeram com que os quatro troncos linguísticos (*cf.* Figura 5) se diversificassem em família e seus falantes se espalhassem pela África, em um percurso que levou ao desaparecimento de algumas línguas e ao contato com tantas outras, em especial o português europeu, impulsionando uma situação que já era complexa, e proporcionando a formação de novas variedades linguísticas, que hoje são identificadas como pidgins e crioulos (Petter, 2015), como tentaremos conceituar e descrever nesta seção.

<sup>13</sup> Segundo Petter e Araújo (2017, p. 38), essas pesquisas tinham como objetivo estudos comparatistas e descritivistas, a citar: os estudos de comparação lexical e gramatical de J. Greenberg, *The Languages of Africa* (1963), em que temos a primeira proposta de classificação das línguas africanas; na pesquisa de campo e descrição de línguas, temos a obra francesa *Enquête et description des langues à tradition orale*, escrita por Bauquiaux e Thomas (1976) e traduzida para o inglês em 1922 – *Studying and Describing Unwritten Languages*.

<sup>14</sup> Esses mentores nem sempre eram africanistas, mas linguistas cujas propostas teóricas impulsionaram o desenvolvimento de estudos em línguas africanas, a saber: Escola anglo-americana, Alemã, Belga, Francesa.



O contexto intercultural e interlinguístico, destacado nas seções anteriores, condensa as discussões em torno da compreensão sobre a formação do Português Brasileiro e suas singularidades e contrastes em relação à variedade europeia, bem como sobre o papel das línguas africanas na variedade brasileira. Como afirmam Avelar e Galves (2014, p. 242, grifos nossos):

Uma das questões que têm permeado os estudos históricos sobre o português brasileiro diz respeito a saber em que medida as múltiplas dinâmicas de contato interlinguístico estabelecidas na formação da sociedade brasileira contribuíram para produzir as características que permitem contrastar suas variedades com as do português europeu. Se, do ponto de vista lexical, não há dúvidas de que o português brasileiro teve uma forte influência de outras línguas [...], a mesma certeza não é consensual no que diz respeito a propriedades gramaticais. Os debates em torno dessa questão têm se dado de forma polarizada, com defensores do que podemos chamar de hipótese da deriva e hipótese do contato ocupando cada um dos lados da polarização – os primeiros argumentam em favor de que todas as marcas gramaticais do português brasileiro foram trazidas de Portugal, não havendo nenhuma propriedade inovadora devida à ação de contatos interlinguísticos; os segundos defendem *que o português brasileiro apresenta características gramaticais que emergiram como consequência do contato, em particular no que diz respeito à constituição das suas variedades populares*.

É sabido que, desde o século XIX, a questão que envolve a formação do português brasileiro e seus sucessivos processos de variação e mudanças linguísticas vem ganhando holofotes no campo do estudo da linguagem. Nesse quadro científico, opiniões se polarizam em duas principais perspectivas teóricas: da deriva e do contato linguístico.

As discussões em torno da origem do Português Brasileiro foram iniciadas pelo filólogo português Adolpho Coelho e continuadas por Serafim da Silva Neto e, mais recentemente, por linguistas tais como Gregory Guy, John Holm, Fernando Tarallo, Dante Lucchesi, Anthony Naro, Marta Scherre, Volker Noll, Mikael Parkvall, só para citar alguns. Assim, apresentaremos alguns pressupostos dos autores supramencionados, no entanto o objetivo desta seção é refletir em prol da hipótese do contato, oferecendo base teórica para análises em torno da variação na concordância nominal de número, uma marca que possivelmente reflete a ação dos contatos interlinguísticos estabelecidos entre os falantes do português e os falantes das línguas africanas. As leituras e resultados já alcançados neste campo de discussão reafirmam que o grande desafio que se coloca à sócio-história da língua portuguesa do/no Brasil é justamente avançar nas pesquisas que envolvem a formação dessa língua. Acreditamos, assim, que o caminho que tem sido percorrido em trabalhos recentes sobre a formação do português brasileiro em que se buscam informações, tais como a constituição da população brasileira, o nível de letramento do seu povo, o contínuo rural/urbano (Bortoni-Ricardo, 2011), as relações sociais que se estabeleceram entre os diversos segmentos da sociedade têm possibilitado desenhar a imagem

do povo brasileiro, afinal, “[...] cada povo que fala hoje o português o modelou, o recriou à sua imagem” (Lourenço, 2001, p. 132).

Iniciamos com as indagações propostas por Alkmim (2007, p. 465-466), que ainda ressoam hoje:

O silêncio sobre a história lingüística de negros e escravos brasileiros causa uma certa perplexidade: como interpretar a ausência de registro histórico? Por que cronistas, viajantes, historiadores, ao longo do período colonial, pouco ou quase nada relataram a respeito dos usos lingüísticos de uma parcela tão significativa da população do Brasil? Devemos ver, nessa atitude, a manifestação de um absoluto desinteresse pela questão, ditada pelo preconceito em relação às línguas africanas e aos usuários africanos da língua portuguesa? Ou não teria havido, efetivamente, nada de excepcional a registrar? [...]. Mas como se deu o contato entre africanos de distintas línguas e culturas? Como foi o processo de aquisição do português pelos escravos? Por que não foi possível manter o uso das línguas africanas ao lado do português?

Como explicar a história de uma língua tão diversa em suas especificidades? Seria mesmo fruto de um contato lingüístico? Ou um processo natural de mudança lingüística? Essas são questões que têm estado presentes nos grandes debates sobre a constituição histórica do português brasileiro e têm impulsionado, desde as últimas décadas do século XX, pesquisas que apresentam as especificidades em torno dessa língua tão diversa.

Pesquisas desse cunho normalmente são conduzidas sob os “olhares” investigativos da Sociolingüística e da Lingüística Histórica, ciências que se debruçam sobre os fenômenos lingüísticos e tentam explicá-los a partir de seus processos sincrônicos e diacrônicos, respectivamente. Através do aparato teórico-metodológico dessas ciências, é possível a realização de diversas pesquisas sobre a sócio-história do português do Brasil, o que tem fortalecido os debates sobre a sua origem. É nessa linha que se insere o nosso trabalho, o qual, ao mesmo tempo que considera fatores externos (sociais) numa tentativa de projeção da imagem da comunidade afro-brasileira estudada, também leva em conta fatores lingüísticos que podem influenciar na escolha do uso variável na concordância nominal de número.

### 2.3.1 Deriva Secular

Quando se fala em deriva, pensa-se no trabalho pioneiro de Anthony Naro, o qual vem apresentando, juntamente com a pesquisadora Marta Scherre, evidências de que o PB é resultado das nossas origens românica e lusitana, considerando que especificidades do PB representam modificações pelas quais o PE passou no Brasil, as quais já seriam previstas no sistema, entendendo que a mudança transparece na frequência dos usos e que a gramática da língua já pressupunha tais variações/mudanças.

No livro *Origens do português brasileiro*, Naro e Scherre (2007, p. 47) apresentam uma descrição geral do modelo de desenvolvimento do Português Popular do Brasil (PPB), afirmando que 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, o qual, no Brasil, encontrou-se com outras forças que reforçavam e expandiam a direção original dessa língua. Essas pressuposições baseiam-se na noção de deriva linguística nos termos de Edward Sapir (1949[1921]), de que a língua se move ao longo do tempo num curso que lhe é próprio, o que implica que as mudanças linguísticas estão, de certa forma, previstas no processo de evolução natural da língua.

Portanto, a hipótese dos autores difere da hipótese da criouliização ou semicriouliização da língua portuguesa em terras brasileiras, uma vez que acreditam que os processos de variação e mudança do PB não advêm do contato com as línguas que vieram para o Brasil com seus povos, ou das línguas ameríndias, que aqui já se encontravam quando chegaram os colonizadores europeus.

Mesmo negando a influência das línguas africanas e indígenas na formação do português popular do Brasil, Naro e Scherre (2007, p. 17) reconhecem a importância desses povos na configuração social e cultural do país: “longe de negar a importância da influência africana e indígena para nossa cultura”. Eis a grande questão: como desvincular língua e sociedade? Estariam os autores abrindo, “reformulando” conceitos já consolidados pela Sociolinguística?

Para amenizar tal contradição, os autores passam a considerar as línguas africanas e indígenas como *força de atração*, as quais juntas se reforçam para produzir o português popular do Brasil (Naro; Scherre, 2007, p. 25). Todas essas questões são minuciosamente explicadas, a partir de argumentações históricas que favorecem a compreensão da proposta de deriva secular, defendida ainda por alguns pesquisadores que se debruçam sobre as origens do português brasileiro.

Sabe-se que, a partir do trabalho de Silva Neto (1986), há referências à fala dos índios e dos brancos, porém Naro e Scherre (2007) ressaltam que, nas pesquisas desse autor, não se encontra nenhuma evidência documental do português falado pelos africanos trazidos para o Brasil, a não ser algumas menções a africanos que não dominavam a língua portuguesa, mas a língua geral tupi ou outras línguas africanas. Nesse sentido, Naro e Scherre (2007, p. 28) afirmam que “a documentação não transmite nem a mais leve suspeita de que a língua portuguesa falada pelos brasileiros descendentes dos primeiros cativos africanos fosse diferente da fala de brasileiros de outras origens étnicas”. Nesse sentido, Mello (2002, p. 344) afirma que

no estudo da mudança linguística, a maior dificuldade que se encontra, ao se tentar avaliar um período remoto na história de uma língua, é a falta de dados associados à variabilidade intrínseca de qualquer comunidade linguística. Portanto, normalmente o que se faz é descartarem-se os idioletos, quando encontrados, e enfocar-se em uma amostragem que se considere representante da comunidade linguística em estudo. Mas não se pode esquecer que a deriva é consequência justamente da escolha, por parte da comunidade linguística, de formas que terão predominância sobre outras dentro de uma ecologia da variabilidade linguística.

Dessa forma, a hipótese da deriva secular, apregoada por Spair e defendidas por Naro e Scherre (2007), apresenta a ideia de que o PB é, na verdade, uma espécie de continuação do português arcaico, com alterações/variações. Assim, acreditam que

[...] o português moderno do Brasil é o resultado natural da deriva secular inerente na língua trazida de Portugal, indubitavelmente exagerada no Brasil pela exuberância do contato de adultos, falantes de línguas das mais diversas origens, e pela nativização desta língua pelas comunidades formadas por estes falantes (Naro; Scherre, 2007, p. 69).

E argumentam em favor da ideia de que

no português do Brasil inexistente influência gramatical específica de qualquer língua africana, ou de língua de qualquer outra proveniência não portuguesa, como também não existe nenhuma forma ou estrutura inteiramente nova criada por um processo geral de simplificação durante a fase de aquisição da língua (Naro; Scherre, 2007, p. 182).

Lucchesi (2012, p. 269) critica essa perspectiva, adota por Naro e Scherre, afirmando que

os fatos aduzidos por Naro e Scherre poderiam, na melhor das hipóteses, servir para refutar a hipótese do contato, não para firmar a hipótese da deriva. Apresentar evidências que neguem a variação que se observa hoje no PB esteja historicamente relacionada ao contato entre línguas não significa demonstrar que essa variação tem a sua origem numa deriva secular. Tal relação não resulta automaticamente de refutação da hipótese do contato linguístico, ela carece igualmente de fundamentação teórica e sustentação empírica [...].

Importante relatar que, mesmo que hoje Naro e Scherre sejam referências ao se tratar da hipótese da Deriva Secular do PB, Joaquim Mattoso Câmara Jr. (1972) já falava nesses termos, quando procurou uma razão estrutural, interna, para explicar o uso brasileiro do *ele* acusativo, na expressão “eu vi ele”. Em seu trabalho, Câmara Jr. (1972, p. 75) exclui a possibilidade de um crioulo de base indígena, porque as línguas indígenas “foram substituídas no intercuro dos índios com os brancos por uma língua única – o chamado Tupi”, restringindo-se aos empréstimos léxicos sua contribuição ao PB. Quanto às línguas africanas, o autor afirma que os escravos negros se adaptaram ao português sob a forma de um falar crioulo, o que impossibilitaria possíveis mudanças fonológicas e gramaticais profundas sem correspondência

com as próprias tendências estruturais da língua portuguesa (*cf.* p. 75-77). Portanto, Câmara Jr. admite o impacto das línguas indígenas e africanas sobre o PB e, embora não admita a existência de um crioulo africano, em nenhum momento afasta a hipótese da mudança natural, ou hipótese da deriva.

Portanto, conclui-se que, na vertente da deriva secular, tenta-se provar a inexistência de traços linguísticos do português falado no Brasil que sejam diretamente ligados à etnia afro-brasileira, o que exclui hipóteses de uma possível língua pidgin ou crioula de base lexical portuguesa, uma vez que estudos, na perspectiva da Deriva, têm evidenciado o fato de que características morfossintáticas e fonológicas do português brasileiro são heranças românicas e portuguesas arcaicas e clássicas, e não modificações advindas das línguas africanas, ou das línguas dos povos ameríndios, como afirmam estudiosos que buscam explicar a origem do PB através do contato linguístico. Em contrapartida, Baxter e Lucchesi (1997), Lucchesi (2001, 2012) acreditam que o contato entre línguas, estabelecido no nosso território durante a colonização, poderia fornecer-lhes uma explicação mais palpável para a formação do PB (Baxter; Lucchesi, 1997; Lucchesi, 2001, 2012).

### 2.3.2 Contato linguístico e sua aplicação ao contexto brasileiro

O contato entre línguas na constituição da história do PB é um tema que tem sido amplamente discutido desde o século XIX, quando já se observava que alguns usos pareciam ser característicos do português falado no Brasil, o que o distanciava do português europeu. Os estudos desde então têm avançado em torno desse assunto, e muito se tem discutido sobre a origem do PB, buscando-se comprovações em todos os níveis de análise linguística: semântico, fonético, morfológico e sintático.

A hipótese do contato, como já é possível inferir a partir de sua própria designação, acredita que as dinâmicas de contato linguístico tiveram um papel fundamental nas especificidades do português brasileiro, frente ao português europeu, especificamente nas variedades do português popular brasileiro. Diante das reflexões de Rosa Virgínia Mattos e Silva (2001, p. 298-299) sobre a formação do português brasileiro, acredita-se que este resulta de vários processos sócio-históricos e linguísticos, marcados pelo multilinguístico/multidialealismo, a saber:

- a. *o português europeu* na sua dialeção diatópica, diastrática, que teria ao longo do período colonial um contingente de 30% da população brasileira; seria esse

*português europeu*, base histórica do *português culto brasileiro* que começaria a elaborar-se a partir da segunda metade do século 18;

b. *as línguas gerais indígenas*, que, plurais e dialetalizadas, poderiam até confundir-se com o *português geral brasileira* nas áreas geográficas delimitáveis em que se difundiram;

c. o *português geral brasileiro*, antecedente histórico do *português popular brasileiro* que, adquirido na oralidade e em situações de aquisição imperfeita, é difundido pelo geral do Brasil sobretudo pela maciça presença africana e dos afro-descendentes que perfizeram uma média de mais de 60% da população por todo o período colonial (Mattos e Silva, 2001, p. 298-299, grifo nosso).

Na perspectiva do contato, temos o filólogo e linguista Serafim da Silva Neto, que já falava nesses termos, ao qual devemos os créditos pelas pesquisas linguísticas e filológicas desenvolvidas sobre o português do Brasil, já que, desde 1960, em artigo publicado no livro *Língua, cultura e civilização*, no trabalho intitulado *Problemas do português da América*, dizia:

A linguagem falada em nossa terra, em virtude de múltiplos fatores, tomou cunho próprio [...] o modo de viver modifica-se e transforma-se. E a língua, instrumento social, foi-se adaptando à nossa sociedade pois, como sabemos, ela caminha lado a lado com a história social. (Silva Neto, 1960, p. 248-258).

Nesse cenário, retomamos e reafirmamos, mais uma vez, a importância de Silva Neto, que, a partir do seu clássico livro *Introdução ao estudo da língua portuguesa no Brasil* (1950), inicia a tentativa de uma reconstrução do passado da língua portuguesa no Brasil e destaca o contato da língua da colonização com a língua dos *aloglotas*<sup>15</sup>.

Como já observado, a partir dos dados de Mussa (1991), o fato de os africanos e afrodescendentes comporem a maior parte do contingente populacional em território brasileiro, entre os séculos XVII e XIX, é inegável, mas as informações em relação ao “nível” e à “forma” da influência que essas línguas e esses povos exerceram na configuração singularizadora do português brasileiro é o grande enigma, a grande questão a ser respondida pelos que acreditam que as especificidades do PB resultam de um contato linguístico.

Muitas são as hipóteses que nos auxiliam na busca de evidências plausíveis em torno da reconstituição histórica do PB. Gregory Guy (2005) acredita na transferência de traços da língua materna (no caso dos africanos) para a segunda língua (o português), sendo que as variedades do PB resultam de um processo de criouliização, hipótese também defendida por Holm (1987). Já Petter (2009) reconhece o protagonismo dos africanos, mas não condiciona as variedades do PB a um processo de criouliização, uma vez que, em linhas gerais, defende que variedades do português constituiriam um *continuum* afro-brasileiro. Nesse viés, Lucchesi (2003; 2009) vem paulatinamente comprovando uma nova vertente que busca evidências de

<sup>15</sup> Termo utilizado por Serafim da Silva Neto (1950) ao se referir às línguas indígenas brasileiras e às línguas africanas trazidas para o Brasil.

que o PB ultrapassa os limites delineados pelos que acreditam em mudanças previsíveis e contínuas inerentes à língua, já que acredita e defende que o contato linguístico resultou em um “processo de transmissão linguística irregular ou croulização do tipo leve, que estaria na base da formação das atuais variedades populares do português do Brasil” (Lucchesi, 2009, p. 71-72).

### 2.3.2.1 Conceituando pidgins e crioulos

Para discutir, especificamente, sobre os conceitos de língua, tomamos como base os pressupostos da Crioulística, ciência que tem focado seus estudos nos processos históricos de contato entre línguas, dos quais são originadas línguas gramaticalmente distintas das envolvidas no processo de contato. Nessa perspectiva, de que a história externa da língua é uma evidência positiva em favor de hipótese crioula, temos trabalhos como os de Coelho<sup>16</sup> (1886), Silva Neto (1950), Guy (1981), dentre outros, que têm atestado e/ou alegado uma origem crioula para o PB.

Nesse viés, temos os trabalhos de Renato Mendonça e Jacques Raimundo (1933 *apud* Bonvini; Petter, 1998) que elaboram os debates sobre criouliização, afirmando que as diferenças ocorrentes no PB se devem à influência das línguas africanas, principalmente o kimbundo e o yorubá (Bonvini; Petter, 1998, p. 1). No que tange às discussões sobre a relevância do contato entre línguas na formação histórica do português do Brasil, Lucchesi (2012, p. 248) afirma que essas questões voltaram à cena na linguística brasileira na década de 1980, com a intervenção de Gregory Guy<sup>17</sup> (1981) e John Holm<sup>18</sup> (1987; 1992).

Guy (1981) e Holm (1987) retomam a discussão sobre a hipótese crioula do PB apoiando-se em dados sócio-históricos, mais precisamente demográficos, e na evidência linguística. Para Guy (1981), a hipótese da criouliização apresenta duas evidências, uma histórica e outra linguística. Em relação à evidência histórico-social, ele acredita que o contexto do Brasil era propício para a formação de um crioulo. Assim, desse ponto de vista, acredita que a criouliização do PB existiu, mas que o sistema linguístico brasileiro está se descriouliizando, e

<sup>16</sup>Adolfo Coelho (1886) foi quem inaugurou os debates sobre criouliização, lançando a ideia da origem crioula no PB.

<sup>17</sup>Guy trabalha com “[...] base em análises variacionistas da fala de indivíduos não escolarizados do Rio de Janeiro, postulando que as condições sociais dos primeiros séculos da colonização do Brasil eram muito favoráveis à criouliização e que o português popular do Brasil seria o resultado de um acelerado processo de descriouliização de um crioulo português que se teria formado no século XVII” (Lucchesi, 2012, p. 249).

<sup>18</sup>Holm “resgatou o conceito de *semicrioulo*, utilizado por Serafim da Silva Neto na década de 1950, defendendo que o português teria passado no Brasil, devido ao contato linguístico, por um processo de *reestruturação parcial*” (Lucchesi, 2012, p. 249)

as forças dessas mudanças linguísticas estão retomando vagarosamente ao seu alvo original. Ainda nessa perspectiva, Guy (1981) acredita que a evidência linguística provém de duas variáveis fonológicas: apagamento do –s e desnasalização de vogais finais; e de duas variáveis morfológicas: concordância no sintagma nominal e no sintagma verbal – o último, objeto de nosso estudo a ser desenvolvido na *Seção 5*.

Para iniciarmos a discussão sobre os conceitos de *pidgin* e crioulo, faz-se necessário entender como surge um *pidgin*. De forma geral, e baseando-se em pesquisas que tratam sobre o tema, pode-se afirmar, ainda que de forma preliminar, que o *pidgin* surge do contato entre línguas mutuamente ininteligíveis e torna-se uma língua auxiliar para comunicação entre indivíduos cujas línguas maternas são diferentes. Porém estudiosos a definem de uma forma mais precisa. Naro e Scherre (2003, p. 286) acreditam que

o pidgin costuma surgir a partir de um contato entre grupos populacionais para propósitos bem definidos e delimitados, tais como trocas comerciais ou trabalho forçado em uma plantação, em um contexto em que as pessoas não dispõem de outro meio verbal comum de comunicação.

Lucchesi (2003, p. 273) também afirma que um *pidgin* surge

quando uma grande população de adultos – em muitos casos falantes de línguas diferenciadas e mutuamente ininteligíveis – é forçada a adquirir uma segunda língua emergencialmente em função de relações comerciais e/ou de sujeição [...].

Tentaremos explicar melhor esse conceito através do processo de formação do português do Brasil. Como já discutido em seções anteriores, o processo de colonização europeia, entre os séculos XVI e XIX, ocasionou várias situações de contato entre povos com diferentes línguas, portanto houve diversas situações propícias ao surgimento de *pidgin* e/ou línguas crioulas. No encontro entre esses povos/comunidades colonizadas, iniciou-se a busca pelo estabelecimento da comunicação entre os escravos e os senhores e entre os próprios escravos, que até então eram falantes de línguas diversas. Diante da necessidade de se comunicar, os falantes trazidos pelos colonizadores de comunidades/línguas diferentes, buscavam uma linguagem comum.

Esse contato interlinguístico tem sido a fonte para a infindável discussão sobre a origem do português brasileiro, até porque muitos são os conceitos adotados pelos pesquisadores, o que torna mais difícil o objetivo, deste subitem, de conceituar *pidgin* e crioulo.

Iniciaremos expondo que uma língua crioula nasce a partir do contato entre comunidades linguísticas diferentes, em meio a uma relação que se estabelece do grupo dominante com o grupo dominado. Antes de tornar-se uma língua crioula, o resultado do contato entre essas etnias será um *pidgin*, um meio de comunicação simplificado.



Nessa interação verbal dos indivíduos dos grupos dominados, ocorre uma negociação em torno das soluções gramaticais empregadas para a formação de frases a partir do vocabulário básico da LA compartilhado por eles, e as estruturas gramaticais do *pidgin* vão se formando, na medida em que sejam consagradas pelo uso, dentro da rede de relações sociais que também se vão formando nessa comunidade emergente. Dessa forma, o *pidgin* vai se expandindo, na medida em que as regras gramaticais empregadas no seu uso são chanceladas dentro de uma rede de relações sociais historicamente determinada. No plano sócio-histórico, os *pidgins* tendem a desenvolver-se nos casos em que os grupos do substrato são mantidos em seu universo cultural de origem e conservam o uso de suas línguas nativas (Lucchesi; Baxter, 2009, p. 105).

Então podemos entender que, a partir do momento que começam a nascer crianças na comunidade, elas passam a ter o *pidgin* que seus pais falam como *input* na fase de aquisição da linguagem; o *pidgin* passa, então, por ação dessa nova geração de falantes, a ser um crioulo, ou seja, o meio de comunicação utilizado pelos adultos naquela comunidade passou a ser uma língua natural para as crianças daquele meio, que o dotam de uma gramática, e assim temos uma língua natural.

Os autores seguem afirmando que a existência de *pidgins* expandidos comprova que a reestruturação gramatical da variedade que se forma na situação de contato pode ocorrer independentemente da sua nativização, o que é tradicionalmente definido como criouliização.

Nesse contexto, Lucchesi (2003, p. 103) aponta que

a noção da criouliização variável foi proposta de maneira explícita por Baker (1982) e Bickerton (1984), que afirmam que a criouliização a partir de situações de contato com maior acesso aos modelos da língua alvo deram origem a crioulos gramaticalmente mais próximos da língua alvo.

Mikael Parkvall, em seu livro *Da África para o Atlântico*, livro traduzido no Brasil, em 2012, por Rodolfo Ilari, afirma que muito se tem discutido a respeito da distinção entre *pidgins* e crioulos, e nas discussões tem sido reconhecido

cada vez mais que a nativização não é uma condição *sine qua nom* da criouliização. Hoje em dia, costuma-se admitir que os *pidgins* podem expandir-se de modo a formar línguas plenas mediante apenas um uso frequente, e foi com base nisso que os ‘*pidgins*’ por exemplo da Nigéria, de Camarões e da Guiné-Bissau foram tratados como crioulos e não como *pidgin* (Parkvall, 2012, p. 33).

Assim, entendemos que os *pidgins* e crioulos se formam a partir de mecanismos radicais de reestruturação da gramática original da língua alvo e de transferência, para esta, de estruturas provenientes das línguas dos segmentos dominados (Lucchesi, 2003, p. 272). Porém Baxter (1994) e Lucchesi (2003, p. 273) afirmam que os contextos sócio-históricos em que ocorre contato entre línguas nem sempre dão margem à implementação de mecanismos mais

radicais de reconfiguração de uma língua. Para ele, as causas da redução na estrutura gramatical da língua de superstrato, advindas do contato com outras línguas, são três:

- a) o difícil acesso dos falantes das outras línguas aos modelos da língua alvo, sobretudo nas situações em que os falantes dessa língua alvo são numericamente muito inferiores aos falantes das outras línguas;
- b) o fato de os falantes dessas outras línguas serem, em sua grande maioria, adultos, não havendo, pois, o acesso aos dispositivos da *faculté du langage*, que atuam naturalmente no processo de aquisição da língua materna;
- c) a ausência de uma ação normatizadora, ou seja, de uma norma ideal que oriente e restrinja o processo de aquisição/nativização, já que esse processo tem como objetivo, fundamentalmente, a comunicação emergencial com os falantes da língua alvo (Lucchesi, 2003, p. 273).

Dessa forma, acredita-se que língua crioula é a evolução do código emergencial implementado no início do contato linguístico, ou seja, é a nativização deste, propiciada pela socialização do código emergencial entre os segmentos dominados. De acordo com Lucchesi (2003, p. 273), este último “vai se convertendo em modelo para a aquisição da língua materna dos descendentes dos falantes das outras línguas”, para cujo *status* ocorre uma expansão gramatical ou por meio da gramaticalização de itens lexicais originários da língua alvo ou por mecanismos gramaticais disponíveis nas línguas de substrato.

No que tange aos critérios linguísticos e estruturais, Naro e Scherre (2003, p. 286) acreditam que

normalmente, o léxico do pidgin se baseia na língua do grupo socialmente dominante, mas, pelo menos no início do processo, as estruturas empregadas podem variar de falante para falante, de acordo com seus conhecimentos linguísticos anteriores e a experiência colhida em suas interações com outros falantes na situação de contato.

Parkvall (2012, p. 43) relata a dificuldade em compreender como funciona o contato linguístico e a reestruturação das funções. Ao discutir as características linguísticas que compõem um crioulo, o autor afirma que deveria ser consensual afirmar que essas características deveriam ter uma linha de princípio, com uma ou mais das quatro origens:

- (i) A língua lexificadora;
- (ii) Os substratos;
- (iii) Os universais da reestruturação;
- (iv) O desenvolvimento independente, incluindo mudanças posteriores à cristalização, quer as motivadas internamente, quer as que resultaram de influência do adstrato (Parkvall, 2012, p. 43-44).

Parkvall (2012, p. 56) ainda expõe que os crioulistas precisam considerar não só a língua lexificadora e o substrato, mas também as propriedades gerais dos *pidgins* e crioulos e as descobertas da linguística tipológica e da linguística histórica. Portanto, o autor propõe

divisões que considera essenciais para a compreensão da reestruturação linguística e os resultados de um contato linguístico. A saber:

- Que um traço linguístico de um P/C seja encarado como uma *conservação da língua lexificadora* se (= se e somente se) estiver presente na língua lexificadora e ausente dos substratos, e se for incomum interlinguisticamente e não geralmente presentes em outros P/Cs não relacionados;
- Que estamos lidando com uma *transferência a partir do substrato* se o traço estiver presente nos substratos, ausente na língua lexificadora e se for translinguisticamente incomum e não geralmente presente em outros P/Cs não relacionados;
- Analogamente, um traço é um *universal de reestruturação* e se está ausente tanto na língua lexificadora quanto nos P/Cs não relacionados;
- *Um caso de desenvolvimento independente*, por fim, caracteriza-se por estar ausente de todos os componentes input, assim como de outros P/Cs não relacionados (Parkvall, 2012, p. 56-57, grifos do autor).

Porém muitos pesquisadores discordam veementemente de que haja o processo de criouliização no caso brasileiro. Dentre esses, Dante Lucchesi (2009, p. 62), em um raciocínio lógico/matemático, diz que “o nível de criouliização é inversamente proporcional à facilidade de acesso aos modelos da língua alvo”, o que implica dizer que, quanto menor o número de falantes da língua dominante, o português europeu, maior será a chance de ocorrer criouliização, mas o autor afirma que esse não é o contexto demográfico do Brasil colônia, uma vez que

[...] a população branca no Brasil nunca foi inferior a 30%, crescendo significativamente no século XIX, quando chega a quase a metade do total. Esse quadro está bastante distante do que deu ensejo às formas mais típicas de criouliização, ocorridas no Caribe. No Haiti e na Jamaica, a proporção de brancos nunca ultrapassou os 10% durante praticamente todo o período da colonização, e o nível de miscigenação entre brancos e negros foi muito mais baixo do que o observado no Brasil. Por tanto, [...]o acesso aos modelos da língua-alvo no Brasil sempre foi maior do que o que se observa nas situações típicas de criouliização (Lucchesi, 2009, p. 62-63, grifos do autor).

Entre esses debates referentes à caracterização do português brasileiro, surge uma nova hipótese que se encontra no rol das discussões sobre a origem do português brasileiro, e reconhece a formação do PB como fruto de um processo de transmissão linguística irregular, em que o contato do PE com diferentes línguas no Brasil, mais diretamente as africanas e indígenas, permitiu o surgimento de uma variedade da língua portuguesa, como discutiremos a seguir, em 2.3.3.

Lucchesi (2012, p. 52) acredita que desenvolvimentos teóricos mais recentes da *crioulística*, o ramo da linguística que focaliza as situações radicais de contato entre línguas, podem ter um papel decisivo para ampliar a compreensão acerca da mudança linguística, desde que devidamente articulados com os achados empíricos advindos da pesquisa de campo sobre

variedades populares de línguas transplantadas no processo histórico da expansão colonial europeia. Dessa forma, as reflexões da crioulística não se circunscreveriam apenas às situações específicas e bem localizadas de pidginização e crioulização, abarcando também as situações de contato massivo que estão presentes na história da grande maioria das línguas do mundo.

Lucchesi (2012, p. 53) ainda afirma que é preciso compreender como os processos de mudança que atuam na pidginização/crioulização podem atuar também em situações de contato cujo resultado não é necessariamente a formação de uma língua pidgin ou crioula. Assim, defenderemos nesta tese que

o debate sobre a hipótese da crioulização prévia: se variedades do português brasileiro foram alguma vez crioulizadas, vem sendo discutido em diversos estudos nas áreas da sociolinguística e crioulística modernas [...]. Hoje em dia, o debate a HCP está numa fase de impasse. Os procedimentos metodológicos tradicionalmente adotados pelos lados por e contra não permitem uma avaliação crítica adequada da hipótese. Para sair deste impasse, *proponho que se adote uma abordagem fatorizada que focaliza as vertentes do português brasileiro que historicamente teriam sido mais propensas à influência de processos de contato de línguas, ou seja, as vertentes rurais africanas e ameríndias* (Baxter, 1995, p. 79-73).

Com esse intuito, o pesquisador Dante Lucchesi tem desenvolvido nos últimos anos o conceito de *transmissão linguística irregular* (Baxter, 1995; Lucchesi, 2003, 2004 e 2008; Lucchesi; Baxter, 2009; Baxter; Lucchesi, 1997) e acredita que, através dessa hipótese, podemos prover a formação histórica do português brasileiro de uma explicação mais satisfatória.

### 2.3.3 Transmissão Linguística Irregular

Através do conceito de *Transmissão Linguística Irregular*, tenta-se delinear os parâmetros sócio-históricos e linguísticos que serviram de base para uma visão sistemática dos processos históricos de mudança induzidos pelo contato entre línguas que não resultam na formação de *pidgins* e crioulos típicos (Lucchesi, 2003, p. 272). Nesse sentido, Lucchesi (2012, p. 254) acredita que a ideia de crioulização leve ou de uma semicrioulização encontra a sua formalização teórica no conceito de transmissão linguística irregular<sup>19</sup>, uma vez que se pretende dar conta dos processos de mudança pelos quais a língua portuguesa teria passado no Brasil em função do contato entre línguas.

<sup>19</sup> Lucchesi (2012, p. 254) afirma que “o conceito de *transmissão linguística irregular* guarda grandes similaridades e alguma complementariedade com o conceito de *reestruturação linguística parcial* de John Holm (2004)”.

Antes de adentrarmos no conceito e caracterização da transmissão linguística irregular, é importante esclarecer que esse é um conceito considerado mais amplo do que o de pidginização e criouliização, uma vez que aborda tanto os processos de mudança resultantes do contato entre línguas que originam uma língua nova – pidgin ou crioulo –, quanto os processos nos quais uma língua sofre alterações decorrentes do contato com outras línguas, sem que essas alterações cheguem a configurar o surgimento de uma nova língua.

Portanto, para dar conta das especificidades do português popular brasileiro, entende-se o conceito de transmissão linguística irregular como algo que designa “[...] os processos históricos de contato massivo e prolongado entre línguas, nos quais a *língua do segmento, que detém o poder político é tomada como modelo ou referência para os demais segmentos*” (Lucchesi, 2003, p. 272, grifo nosso). A partir desse conceito preliminar, poderemos iniciar as discussões em torno da caracterização do processo de transmissão linguística irregular do PB.

Talvez surjam dúvidas sobre como tal processo pode produzir variações na língua receptora. Nesse sentido, Lucchesi e Baxter (2009, p. 101) nos explicam que

tal processo de nativização da língua dominante ocorre de maneira irregular no sentido de que os dados linguísticos primários de que as crianças que nascem nessas situações 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. Tal processo diferencia-se da situação de transmissão geracional normal das línguas humanas, em que as crianças dispõem de dados linguísticos mais completos fornecidos pela língua materna dos seus pais.

Nesse contexto, há duas possibilidades que podem ser geradas por uma transmissão linguística irregular: a formação de uma língua crioula ou uma nova variedade da língua de superstrato, que é representado pelos processos de variação e mudança induzidos pelo contato entre línguas (Lucchesi; Baxter, 2009).

No Brasil, a língua portuguesa representava superioridade econômica e política, era o caminho para uma ascensão social. Com esse intuito, muitos colonizados viram na língua de “prestígio” uma chance para ascensão social, porém adquirir uma nova língua, naquele contexto social e multicultural, não seria tão simples. Dessa forma, a língua portuguesa chega até os colonizados/escravizados já apresentando algumas alterações morfossintáticas, resultando numa variedade da língua alvo, como atestam alguns trabalhos de Lucchesi (2001; 2003; 2009), que afirma que

prolongando-se a situação de contato, essa variedade segunda da língua alvo, por ser socialmente mais viável (em função de sua relação com os segmentos política e

economicamente dominantes e de maior prestígio cultural), vai progressivamente assumindo novas funções na rede de interação linguística, ao tempo em que se vai convertendo em modelo para a aquisição da língua materna dos descendentes dos falantes das outras línguas (Lucchesi, 2003, p. 273).

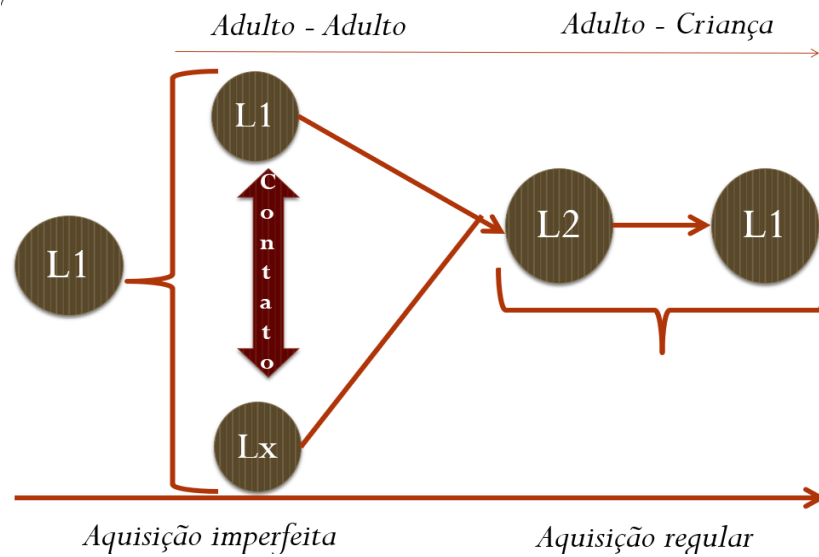
Lucchesi (2009) afirma que, na história sociolinguística do Brasil, não ocorreram processos que possam ser caracterizados como crioulização. Portanto, tal afirmação dá margem a entender que o contato entre línguas, no período colonial, deu origem a uma variedade de língua a ser explicada pelo processo de transmissão linguística irregular. Dessa forma, Lucchesi (2009, p. 71) explica que

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.

O conceito de transmissão linguística irregular tomado por Lucchesi e Baxter (2009, p. 101) busca analisar de forma ampla 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”. Portanto, torna-se inevitável retroceder ao contexto sociolinguístico no período colonial para entender como se configura essa situação de contato e como o PB pode ter sido originado da transmissão linguística irregular.

Como já retratado nas seções anteriores, o Brasil se apresenta inserido em um contexto plurilinguístico, em que havia a língua do dominador, o português europeu, denominada língua de *superstrato* ou *língua-alvo*, e as muitas línguas dos dominados, denominadas línguas de *substrato*. Como em todo processo de colonização, a língua de superstrato se impõe às demais línguas, de modo que “[...] os falantes das outras línguas, em sua maioria adultos, são forçados a adquiri-la em condições bastante adversas de aprendizado, em função de sua sujeição e marginalização” (Lucchesi; Baxter, 2009, p. 101). Os autores ainda afirmam que essas aquisições “imperfeitas” de segunda língua, e nas condições em que foram adquiridas, acabam por fornecer modelos de aquisição da língua materna para as novas gerações de falantes. Isso se configura no conceito do que Lucchesi (2003; 2009) denominará de transmissão linguística irregular. Vejamos uma tentativa de representação desse conceito:

**Figura 6** – Representação para o conceito de Transmissão Linguística Irregular



Fonte: Elaborado pela autora.

A variedade de língua alvo que se forma de contato plurilinguístico apresenta simplificações em sua estrutura gramatical. Lucchesi (2009) acredita que da língua alvo se mantêm os elementos essenciais necessários ao preenchimento das funções comunicativas básicas; no mais, apresentam-se altos índices de variação em relação ao modelo de língua alvo.

Em linhas gerais, Lucchesi e Baxter (2009) postulam duas fases presentes no processo de transmissão linguística irregular. Vejamos:

- (i) Fase inicial: 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;
- (i) Fases seguintes: manutenção do quadro de variação, em que se observa a concorrência entre as formas gramaticais reintroduzidas a partir da influência da LA e potenciais processos de gramaticalização que emergem na heterogeneidade da fala; variação nas frequências de uso dos parâmetros sintáticos (Lucchesi; Baxter, 2009, p. 123)

Lucchesi (2003, p. 276) apresenta três características fundamentais do processo de transmissão linguística irregular:

- (i) Perda, ou variação no uso, de morfologia flexional e palavras gramaticais;
- (ii) Alteração dos valores dos parâmetros sintáticos em função de valores não marcados, que não implicam, entre outras coisas, movimentos aparentes na estruturação da sentença; e
- (iii) Gramaticalização de itens lexicais para preencher as lacunas na estrutura linguística.

As mais recentes pesquisas no campo da Transmissão Linguística Irregular têm demonstrado que esse é um amplo conceito que tem sido capaz de abarcar e interpretar muitos processos históricos em que o português sofreu alterações/variações. Dessa forma, no campo da ciência da linguagem, torna-se essencial buscar evidências/fatos que venham a ratificar sua importância no contexto de origem do português brasileiro. Portanto, as mostras de fala da comunidade rural afro-brasileira Volta do Angico serão aqui interpretadas a partir da perspectiva sócio-histórica, e a constituição de um *corpus* oral representativo de comunidades rurais afro-brasileiras do Território de Identidade de Irecê, desenvolvido no âmbito do Projeto ELIHS (*cf.* seção 4.2), será organizada a partir da sugestão proposta por Mattos e Silva, como uma via possível para recuperar o PB popular.



### 3 SOCIOLINGUÍSTICA: PRESSUPOSTO TEÓRICO PARA UM RETRATO DOS PROCESSOS DE VARIAÇÃO E MUDANÇA NO PORTUGUÊS BRASILEIRO

*“Toda a teoria deve ser feita para poder ser posta em prática e toda a prática deve obedecer a uma teoria. Só os espíritos superficiais desligam a teoria da prática, não olhando a que a teoria não é senão uma teoria da prática, e a prática não é senão a prática de uma teoria. Na vida superior a teoria e a prática completam-se. Foram feitas uma para a outra.”*  
*Fernando Pessoa*

Os estudos linguísticos que operam com as investigações dos fatos da língua podem demonstrar, a depender da orientação teórica que assumem, que a linguagem se constitui ferramenta da interação humana em suas atividades socialmente convencionalizadas, o que impõe a ideia de que os interlocutores também agem no processo de troca e constroem os sentidos de acordo com o seu propósito no contexto de ação pela linguagem.

Até adquirir essas concepções que hoje se têm da própria língua e dos processos subjacentes à atividade de uso da linguagem para fins comunicativos, a ciência linguística se viu envolvida em sucessivas etapas de aprimoramento investigativo e avanços teóricos que possibilitaram aos linguistas reverem perspectivas associadas a uma visão puramente imanente. Portanto, antes de apresentarmos as abordagens teóricas que norteiam esta tese, acreditamos ser importante situar, sem pretensões de historicizar, o processo de aperfeiçoamento dos estudos linguísticos e sua ancoragem na visão de língua como instrumento de comunicação e interação social, que recairá na reflexão sobre o uso e na teoria da Sociolinguística Variacionista.

Assim, discutiremos sucintamente alguns marcos históricos da linguística, compreendendo que desde muito tempo, antes mesmo de se consolidar como ciência, já fornecia ao homem importantes informações acerca da língua e da linguagem usada pelos seres humanos. Iniciaremos, desta forma, com a apresentação de um breve panorama da ciência linguística, apontando para alguns de seus principais avanços até o desencadeamento e consolidação do modelo Sociolinguístico Quantitativo, que apresenta o método de trabalho que fundamentará a pesquisa que ora se apresenta. Este caminhar favorecerá a compreensão em relação as escolhas teórico-metodológicas adotadas para o trabalho empírico apresentado, afinal, retomando nossa epígrafe, *“toda a teoria deve ser feita para poder ser posta em prática e toda a prática deve obedecer a uma teoria”* (Fernando Pessoa).

### 3.1 Linguagem e Linguística: considerações gerais

O descrever sobre o estabelecimento dos estudos linguísticos diante dos processos de variação e da mudança permite destacar algumas concepções de língua a partir de distintos modelos teórico-metodológicos, desvelando questões epistemológicas relacionadas à ciência linguística. Assim, iniciamos afirmando que a história nos permite afirmar que o atual quadro teórico que existe no campo da Linguística é, sem dúvidas, o resultado de sucessivas etapas de revisões teóricas, reformulações de hipóteses sobre os fatos da língua e atualizações constantes dos pontos de vista acerca da própria linguagem.

Ao se traçar um panorama dessas etapas fica evidente que ao longo do tempo os estudos da linguagem foram se aperfeiçoando até adquirirem o status de ciência. Dessa aquisição resulta o aprimoramento das investigações que se têm sobre a linguagem, bem como o desencadeamento de modelos teóricos diversos que dispõem de variadas percepções acerca da língua e dos campos de análise da linguagem<sup>20</sup>.

Iniciaremos refletindo, brevemente, sobre o fato da linguagem, ao longo de sua história, sempre ter ocupado o lugar de objeto de investigação do homem, mesmo em épocas em que não se pensava num campo do conhecimento humano responsável por descrever e analisar o funcionamento e uso da língua, conferindo-lhe o status de ciência. O primeiro registro histórico que se tem é o da Índia Antiga, que, por razões meramente religiosas, desenvolveu uma investigação da fonética articulatória, com a intenção de “manter viva a pronúncia correta dos textos religiosos ancestrais” (Weedwood, 2002, p. 17), enquanto isso, na Grécia, Platão foi o responsável pelo estudo da relação entre conceito e significado, ou seja, buscava compreender até que ponto a palavra se relacionava com o seu significado. Conforme Weedwood (2002), as inquietações geradoras das investigações do filósofo grego partiam do questionamento da possível existência de vínculo direto e essencial entre a língua e a realidade. A pretensão da filosofia grega era esclarecer em que plano se efetuava a conexão entre as palavras e aquilo que denotavam, se por meio da *phýsis*— “a natureza, o poder inexorável que governa o mundo

---

<sup>20</sup> Conferir o livro “Conversas com Linguistas”, organizado por Cortez e Xavier (2003), em que apresentam como as concepções em relação a ciência da linguagem no Brasil vão se modelando e aperfeiçoando, através de entrevista com experientes estudiosos da linguagem - Maria Bernadete N. Abaurre; Eleonora C. Albano; José Borges Neto; Ataliba de Castilho; Carlos Alberto Faraco; José Luiz Fiorin; João Wanderley Geraldi; Francisco C. Gomes de Matos; Rodolfo Ilari; Mary Kato; Ingedore G. Villaça Koch; Luiz Antonio Marcuschi; Maria Cecília Mollica; Diana Luz P. de Barros; Sírio Possenti; Kanavilill Rajagopalan; Margarida Salomão e Carlos Vogt.

visível” (p. 24) – ou por meio da *nómos* (ou *thésis*) – “a crença, o costume ou lei instituída por ação divina ou humana” (p. 24).

Ao longo da Idade Média, a língua latina já havia se expandido por muitas regiões da Europa. Com o crescente avanço do latim fez-se necessário uma compilação que reunisse paradigmas de nomes e verbos do vocabulário cristão e exemplificasse formas morfológicas, tais como sufixo, derivacional, gênero e terminação nominal. Mais tarde, todo esse material foi fundido à *Ars Minor* de Donato, o que permitiu o ensino de conceitos gramaticais e seus paradigmas (Weedwood, 2002). Assiste-se, então, “à passagem de uma gramática primordialmente semântica e taxionômica para uma gramática descritiva” (Weedwood, 2002, p. 51). Como afirma Saussure (2006 [1916], p. 7, grifo nosso),

Começou-se por fazer o que se chamava de ‘Gramática’. Esse estudo, inaugurado pelos gregos, e continuado principalmente pelos franceses, é baseado na lógica e está desprovido de qualquer visão científica e desinteressada da própria língua; *visa unicamente a formular regras para distinguir as formas corretas das incorretas*; é uma disciplina normativa, muito afastada da pura observação e cujo ponto de vista é forçosamente estreito.

É fato que o percurso traçado pela linguística contribuiu de modo significativo para a consolidação de descobertas que garantiram ao campo dos estudos da linguagem o status de ciência autônoma. Segundo Faraco (2005 [1950], p. 132), deve-se a William Jones<sup>21</sup> o marco simbólico do início da linguística como ciência, que no final do século XVIII, em 1786, demonstrou que a semelhança constatada entre o sânscrito, o grego e o latim, levantava a hipótese de que havia uma origem comum a essas três línguas<sup>22</sup>. Com esse evento, ganha fôlego na Europa um movimento de estudos comparativos e históricos e sua respectiva contribuição para a construção da ideia de imanência no âmbito da língua (Faraco, 2005 [1950], p. 132-133).

Com a descoberta de William Jones da origem comum, o século XIX vai ser marcado por métodos investigativos baseados no método comparativo. É, pois, a partir de trabalhos de Franz Boop<sup>23</sup> que o método comparativo adquire fundamento empírico no trato do parentesco

<sup>21</sup> William Jones (1746- 1794), cidadão inglês que, dentro da administração colonial britânica da Índia, exercia a função de juiz em Calcutá. Apresentou uma comunicação à Sociedade Asiática de Bengala, destacando as inúmeras semelhanças entre o sânscrito, o latim e o grego (Cf. Faraco, 2005[1950], p. 132).

<sup>22</sup> Além de William Jones, Franz Bopp, Jacob Grimm, Augusto Schleicher, entre outros, foram os responsáveis pelo início do estudo histórico das línguas com base no método comparativo. Os gramáticos histórico-comparatistas entendiam a língua como um sistema homogêneo (como, aliás, de maneira geral, ela vinha sendo vista desde a Antiguidade) e a mudança como relacionada apenas a fatores linguísticos.

<sup>23</sup> Franz Boop (1791-1867) foi formado pela Escola de Estudos Orientais e teve significativa importância, ao lado de Friedrich Schlegel (1772-189), por desenvolver a chamada *gramática comparativa*, que desenvolvia estudos comparativos da língua sânscrita em comparação com a língua grega, latina, persa e germânica. Em suas obras, Franz Boop “demonstrou, pela comparação detalhada da morfologia verbal de cada uma dessas línguas, as

entre línguas. Assiste-se, mais adiante, à criação de áreas especializadas com a pesquisa comparativa. Conforme se lê em Faraco (2005 [1950], p. 136-137), “nessa linha, destaca-se principalmente o desenvolvimento da chamada filologia (ou linguística) românica, nome que se deu ao estudo histórico-comparativo das línguas oriundas do latim [...]”. No entanto, “tal escola, porém [...] não chegou a constituir a verdadeira ciência da Linguística. Jamais se preocupou em determinar a natureza do seu objeto de estudo. Ora, sem essa operação elementar, uma ciência é incapaz de estabelecer um método para si própria” (Saussure, 2006 [1916], p. 10).

Com o aparecimento dos neogramáticos, da Universidade de Leipzig, na Alemanha, em fins do século XIX, “os pressupostos tradicionais da prática histórico-comparativa” passaram a ser questionados e se estabelece uma nova metodologia e uma série de postulados teóricos para a interpretação da mudança linguística. Para essa nova geração de linguistas, era preciso estabelecer “uma orientação metodológica diferente e de um conjunto de postulados teóricos para a interpretação da mudança linguística” (Faraco, 2005 [1950], p. 139). Segundo defendiam os neogramáticos, o interesse pela investigação linguística deveria ser a natureza da mudança, fazendo oposição aos estudos histórico-comparativos, cujo objetivo era desvendar o parentesco entre as línguas. Assim, os neogramáticos postulavam que a pesquisa linguística investigasse os mecanismos da mudança pelos quais sofriam as línguas. Essa visão norteou as bases dos estudos que se estenderam até as primeiras décadas do século XX.

Naturalmente, como todos os modelos teóricos antecedentes, as formulações dos neogramáticos foram alvo de críticas e questionamentos de outros linguistas que não concordavam com as ideias defendidas pelos neogramáticos. A crítica que se faz aos defensores da teoria da mudança fundamenta-se no fato de o arcabouço teórico do método investigativo da mudança linguística, empreendido pelos neogramáticos, não levar em consideração, “na compreensão dos fenômenos da história, as relações entre língua e sociedade, relações que os estudos de sociolinguística têm mostrado serem particularmente relevantes para se entender a mudança linguística” (Faraco, 2005 [1950], p. 149).

Importante frisar que se deve ao linguista austríaco Hugo Schuchardt<sup>24</sup> a consideração

---

correspondências sistemáticas que havia entre elas, fundamento para se revelar empiricamente seu efetivo parentesco” (Faraco, 2005 [1950], p. 133-134).

<sup>24</sup> Hugo Schuchardt (1842-1927), austríaco, um dos principais linguistas que criticam os postulados dos neogramáticos. Embora ainda tivesse o falante individual como ponto de referência – concepção subjetivista da língua – Schuchardt ao se opor ao conceito de lei fonética, chamou a atenção para a imensa gama de variedades

de um quadro heterogêneo, em que o contexto cultural e social da língua atua como condicionantes da variação. Defendia-se a existência das variedades de fala que eram fruto do condicionamento de fatores relacionados ao gênero, à idade e ao nível de escolaridade dos falantes. Essa nova visão contribuiu para que a língua fosse entendida a partir do contexto social em que estava inserida, sendo importante levar em consideração a realidade linguística dos falantes que faziam uso da língua para a comunicação e atuação em sociedade, “mais do que isso, ele mostrou como essas variedades se influenciam mutuamente, como as línguas em contacto — quer pela proximidade geográfica, quer em decorrência de invasões, conquistas e intercruzamentos étnicos e culturais — também se influenciam mutuamente” (Faraco, 2005 [1950], p. 151). Este quadro heterogêneo apresentado por Schuchardt permitiu compreender melhor os processos de mudanças linguísticas que, no decorrer do século XX, desdobrar-se-á em concepções de que o contexto histórico, social e cultural da língua é condicionante básico da variação e, dentro dela, da mudança, abrindo caminhos para os estudos dialectológicos e, mais recentemente, sociolinguísticos<sup>25</sup>, que fundamenta teórica e metodologicamente esta pesquisa.

Em meados do século XX um fato histórico marca a consolidação da Linguística enquanto ciência autônoma que estuda a linguagem: a publicação póstuma, em 1916, do *Cours de Linguistique Générale*<sup>26</sup>, escrito por Chales Bally e Albert Sechehaye, através de notas das aulas do mestre genebrino Ferdinand de Saussure (1857-1913). Nestes escritos se estabelece, portanto, que a linguística se constitui ciência a partir de três fases sucessivas pelas quais passou o estudo dos fatos da língua, a saber:

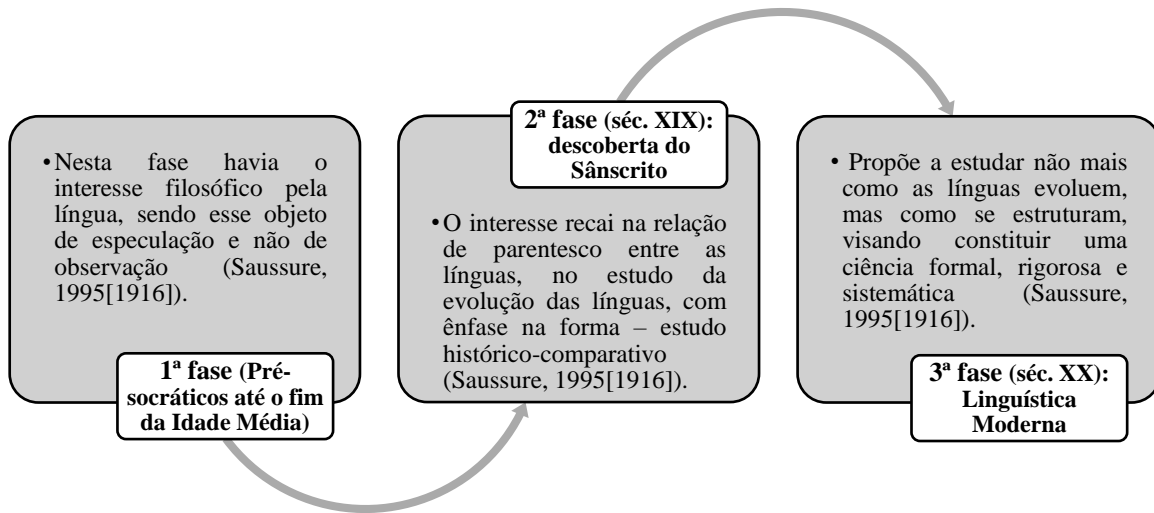
---

de fala existente numa comunidade qualquer, variedades essas condicionadas por fatores como o sexo, a idade, o nível de escolaridade do falante (cf. Faraco, 2005 [1950], p. 151). Mattos e Silva (2008b, p. 42, grifo da autora) afirma que “[...] é Hugo Schuchardt que abre caminhos para uma nova orientação – o estudo da variação linguística no espaço e sua compreensão para as mudanças no tempo; além disso, é o mesmo Schuchardt que vai introduzir nas preocupações dos oitocentistas o fenômeno do contacto entre línguas. Seu livro seminal é publicado em 2ª edição em 1928, com o título de Brevier: *Ein Vademecum der allgemeinen Sprachwissenschaft*.”

<sup>25</sup> Cabe frisar que de Dinah Callou (2009), em capítulo intitulado *De Fernão de Oliveira e da (Socio)linguística*, afirma que fundamentos da atual linguística de base empírica – representada pela Sócio e pela Dialectologia – encontram-se em Fernão de Oliveira, primeiro gramático da língua portuguesa. Portanto, o papel relevante, no âmbito dos estudos linguísticos contemporâneos, que assume a Gramática de Fernão de Oliveira, não advém apenas do fato de ser a primeira gramática da língua portuguesa, mas também do de trazer à tona, no século XVI, fundamentos que constituem o carro-chefe da sociolinguística laboviana, como os da variação de gênero e faixa etária, além da variação diatópica, premissa da Dialectologia.

<sup>26</sup> As próximas referências farão alusão à versão traduzida, *Curso de Linguística Geral* (1995).

**Figura 7** – Fases dos estudos da linguagem



Fonte: Elaborada pela autora.

É no início do século XX, com a divulgação dos escritos de Ferdinand Saussure, no *Curso de Linguística Geral*, que a investigação sobre a linguagem – a Linguística – passa a ser reconhecida como estudo científico, nos termos expostos por Martelotta (2011, p. 20):

O estatuto científico da linguística deve-se, portanto, à observância de certos requisitos que caracterizam as ciências de um modo geral. Em primeiro lugar, a linguística tem um objeto de estudo próprio: a capacidade da linguagem, que é observada a partir dos enunciados falados e escritos. Esses enunciados são investigados e descritos à luz de princípios teóricos e de acordo com uma terminologia específica e apropriada. A universalidade desses princípios teóricos é testada através da análise de enunciados em várias línguas.

Em segundo lugar, a linguística tende a ser empírica, e não especulativa ou intuitiva, ou seja, tende a basear suas descobertas em métodos rígidos de observação. Ou seja, a maioria dos modelos linguísticos contemporâneos trabalha com dados publicamente verificáveis por meio de observações e experiências.

Temos então o marco inicial da linguística contemporânea. Como afirma Lucchesi (2004, p. 39) o *Curso* caracteriza-se como a pedra angular do estruturalismo através da concepção axial de língua como sistema de signos, constituindo-a enquanto legítimo objeto de estudo da linguística. Neste sentido, como aponta Borges Neto (2010, p. 2), “constituir uma teoria qualquer supõe fazer um recorte no objeto observacional e em ‘organizar’ essa porção do mundo a partir de noções teóricas”. Ainda nos apropriando de suas palavras, cedida em entrevista à Revista virtual de estudos da linguagem, comungamos com a analogia que propõe:

Creio que as tentativas de enxergar a linguagem e suas partes como um grande lego podem ser substituídas pela *metáfora da sinfonia*, em que as várias partituras (dos violinos, do oboé, dos contrabaixos, das trompas, etc.), não sendo ‘complementares’ já que se sobrepõem a todo momento, permitem um conjunto harmonioso. Às vezes os violinos se calam enquanto os metais são acionados; às vezes o fagote faz um solo; outras vezes todos os instrumentos tocam simultaneamente; e esses movimentos, no seu conjunto, constituem o todo harmônico (Borges Neto, 2010, p. 4, grifos do autor).

Dito isso, faz-se importante iniciarmos considerando a concepção de língua assumida pelo estruturalismo europeu, também identificado com estruturalismo saussuriano, na tentativa de conhecermos de forma mais aproximativa seu aparato teórico-epistemológico, para que seja possível compreender as divergências epistemológicas existentes entre esta corrente linguística e a Sociolinguística (*cf.* seção 3.2)

Para discutir a natureza do signo linguístico, Saussure mostra que a visão simplista de que a língua é apenas “uma nomenclatura” ou “uma lista de termos que correspondem a outras tantas coisas” é criticável e insuficiente para dizer se a palavra é de natureza vocal ou psíquica. Nesse sentido, o mestre genebrino propõe parâmetros dicotômicos (língua e fala, sincronia e diacronia, significante e significado, paradigma e sintagma) para a nova ciência e estabelece que “a unidade linguística é uma coisa dupla, construída da união de dois termos” (Saussure, 1995 [1916], p. 79).

Para Saussure (1995[1916], p. 17), a *langue* (língua) é um objeto unificado e suscetível de classificação “é um produto social da faculdade da linguagem e um conjunto de convenções necessárias, adaptadas pelo corpo social para permitir o exercício dessa faculdade nos indivíduos”. A *langue* é parte social externa ao falante, enquanto que a *parole* (fala) é entendida como ato individual, correspondendo a atuação do falante, sua emissão e recepção. Com essa dicotomia, o estruturalismo reserva a fala os aspectos inerentes a heterogeneidade, característica das manifestações linguísticas<sup>27</sup>.

Nesse sentido, a língua é considerada uma estrutura constituída por uma rede de elementos, em que cada elemento tem um valor funcional determinado. Ao estabelecer a dicotomia *langue* e *parole*, Saussure estava construindo suas perspectivas de investigação. Ele escolheu a *langue*, pois defendia que o objeto de estudo da linguagem deveria ser coletivo e organizado, não individual e homogêneo.

Para Saussure, cabe ao linguista considerar tanto as relações do sistema linguístico entre si quanto às transformações a que o sistema se sujeita ao longo do tempo. Nesse sentido,

---

<sup>27</sup> Conferir capítulo intitulado *A sistematicidade da mudança: do Estruturalismo à Sociolinguística* (2004), em que Lucchesi discorre sobre o percurso da linguística, a partir da concepção de língua do modelo teórico estruturalista, demonstrando criteriosamente as perspectivas formadoras responsáveis pelo atual panorama da ciência da linguagem.

se delineava a oposição entre uma Linguística sincrônica e uma Linguística diacrônica, ao compreender que o que “é sincrônico tudo que se relacione com o aspecto estático da nossa ciência, diacrônico tudo que diz respeito às evoluções” (Saussure, 1995 [1916], p. 96). Segundo essa distinção, os fenômenos do objeto linguístico devem ser descritos a partir de um estado de língua, daí a designação de sincronia, e também por meio de uma fase de evolução, o que explica a diacronia.

Considerando que os estudos que antecederam o estruturalismo saussuriano adotavam a perspectiva diacrônica, Saussure escolhe a sincronia e busca dar amplo espaço à observação da língua em seu estado no tempo em que se prestava a investigação, pois havia o entendimento de que “o aspecto sincrônico prevalece sobre o outro, pois, para a massa falante, ele constitui a verdadeira e única realidade. [...] na perspectiva diacrônica, não é mais a língua que se percebe, mas uma série de acontecimentos que a modificam” (Saussure, 1995[1916], p. 106). Por defender a imanência da língua como objeto de estudo, Saussure a concebe como sistema de signos, em que pode ser estudada por si mesma, cuja característica principal seria a homogeneidade.

Muitos são os questionamentos em relação as escolhas de Saussure<sup>28</sup>, no entanto, compreendendo como um marco inicial da Linguística Contemporânea, em que se buscava garantir a autonomia e cientificidade frente às demais ciências humanas, é compreensível a rigorosidade na delimitação do objeto de estudo, compreendendo-o em sua dimensão sistêmica e estável.

A respeito das inovações que se consolidam a partir da obra saussuriana, Faraco (2005 [1950], p. 28) aponta que:

[...] seu ovo de Colombo foi não só mostrar que a língua poderia (e deveria) ser tratada exclusivamente como uma forma (livre das suas substâncias), mas principalmente como esta forma se constituía, isto é, pelo jogo sistêmico de relações de oposição – funcionando este jogo de tal modo que não é num sistema linguístico senão por uma teia de relações de oposição. E, por outro lado, nada interessa numa tal perspectiva sistêmica salvo o puramente imanente.

A língua, assim entendida, passa a revelar que a forma linguística é uma entidade que opera funções e que a estrutura do sistema oferece bases para estudos que se voltam para a imanência do objeto linguístico. Sendo assim, enquanto o formalismo projetava a língua como um sistema estruturado, assumindo – o como homogêneo e invariável, observações empíricas apresentava um contexto linguístico extremamente variável. Como expõe Lucchesi (2004, p.

---

<sup>28</sup>Cf. Lucchesi (2004), Castelar de Carvalho (2003 [1997]), Bouquet (2000, [1997]), entre outros.



22-23, grifo do autor), é claro a contradição entre a concepção de língua como um sistema e a questão da mudança linguística:

A visão da sistematicidade da língua remete à questão do funcionamento das línguas enquanto instrumento privilegiado da comunicação humana, a sua condição de código, que, para cumprir suas funções, deve ser estruturado. É no modo estruturante da linguagem que se deve, portanto, buscar a explicação para o fascinante mecanismo provedor de informação referencial que é a língua. A questão da funcionalidade da língua levanta a necessidade de se fixar o sistema linguístico para que se possam analisar os mecanismos do seu funcionamento.

Por outro lado, a questão da mudança levanta uma pergunta igualmente aliciadora: *o que faz um sistema que funciona tão bem mudar?*

Lucchesi (2004, p. 23, grifos do autor) aponta um paradoxo e afirma que “a *contradição entre mudança e sistema* assenta sobre suas dimensões fundamentais e antitéticas do fenômeno linguístico: a sua *dimensão estrutural* e a sua *dimensão sócio-histórica*” e que a tensão que emana deste binômio se fez presente ao longo de toda a história do estudo da linguagem. O autor ainda afirma que o estruturalismo sincrônico procurou resolver essa contradição presente entre sistema e mudança através de um aparato teórico-metodológico em que faz alusão a existência de fatores externos ao fenômeno linguístico.

No período entre 1920 e 1960, notadamente, o estruturalismo é o modelo teórico hegemônico dentro da Linguística,

o surgimento e a ascensão do estruturalismo explicam-se pelo fato de esse modelo ser capaz de superar contradições que então se colocavam no desenvolvimento da teoria linguística, com base na nova concepção do objeto de estudo que ele apresenta. Por outro lado, ao tempo em que promovia sínteses para a superação das contradições com que se debatia a teoria linguística, o estruturalismo também engendrava novas contradições, que, no plano da história interna da linguística, podem explicar em grande medida o seu desenvolvimento e a sua própria superação enquanto modelo teórico hegemônico (Lucchesi, 2004, p. 24).

Assim, a partir de meados do século XX, surgissem propostas teóricas que disputam a hegemonia do estruturalismo, sucedendo este modelo teórico ao apresentarem formulações que permitem a superação das contradições teóricas, pontualmente a contradição entre sistema e mudança. As novas perspectivas linguísticas passam a integrar análises estruturais a fatores sócio-históricos, propostas essas consolidadas, de forma mais sistematizada, no âmbito da teoria sociolinguística variacionista, que assume, em termos gerais, que a variabilidade é inerente ao sistema linguístico, cujo funcionamento é afetado por fatores de ordem interna e externos à língua.

Assim, não podemos deixar de frisar que é no âmbito do funcionalismo linguístico que esses fatores externos começam a ser considerados nos processos de investigação da língua. Ao

se adotar a concepção de que a língua é instrumento de comunicação dos seus usuários, o funcionalismo preconiza, então, que a construção de significados é efetuada tanto pela operação dos recursos formais da língua, que se manifestam através dos textos falados e escritos quanto pela atuação de fatores que condicionam os recursos linguísticos à organização prevista e desejada pelo usuário. Isso significa que a interação verbal é resultado de intervenções de ordem do sistema da língua, bem como de ordem dos sentidos extralinguísticos que se acomodam na caracterização dos usuários, na delimitação de seus propósitos comunicativos e na sua habilidade de manifestar-se textualmente (falando ou escrevendo).

Antes de começarmos a tecer considerações acerca do modelo teórico funcionalista, apoiamo-nos nas palavras introdutórias da *Gramática Funcional* (Neves, 1997) ao afirmar que caracterizar o funcionalismo não constitui uma tarefa fácil, visto que

os rótulos que se conferem aos estudos ditos ‘funcionalistas’ mais representativos geralmente se ligam diretamente aos nomes dos estudiosos que os desenvolveram, não a características definidoras da corrente teórica em que eles se colocam. Prideaux (1994) afirma que provavelmente existem tantas versões do funcionalismo quantos linguistas que se chamam *funcionalistas*, denominação que abrange desde os que simplesmente rejeitam o formalismo até os que criam uma teoria. A verdade é que, dentro do que vem sendo denominado – ou autodenominado – ‘funcionalismo’, existem modelos muitos diferentes (Neves 1997, p. 01, grifos da autora).

A autora segue afirmando que mesmo diante de distintas peculiaridades é possível notar similaridades que unem os modelos e caracterizam a visão funcionalista da linguagem, podendo destacar a ideia de que a língua não pode ser descrita como um sistema autônomo, já que o sistema linguístico só pode ser entendido a partir de noções como cognição e comunicação, processamento mental, interação social e cultura, mudança e variação (Neves, 1997, p. 3). É neste ponto das similaridades que nos apoiamos para escrever, sucintamente, sobre o funcionalismo, a qual desvelará importantes concepções que também nortearão os pressupostos teóricos da Sociolinguística.

Regressando as discussões suscitadas até o momento é possível notar que todo o empreendimento dos estudos linguísticos até o século XIX preparou o espaço que Saussure necessitava para explorar e garantir o caráter científico da Linguística, no século XX. No entanto, cabe ainda ressaltar que ao definir a língua como sistema de signos, o estruturalismo saussuriano ganha novos adeptos e novos grupos que se constituíam em torno da investigação linguística, a exemplo do Círculo Linguístico de Praga, em 1928, que adota o termo *sistema*. O estruturalismo dos linguistas praguenses foi influenciado, também, pelas ideias do psicólogo

alemão Karl Bühler<sup>29</sup>, as quais contribuíram para a consideração do aspecto funcionalista. A partir das propostas da Escola Linguística de Praga o termo *função*<sup>30</sup> passou a ser aplicado para designar o “papel desempenhado por um elemento estrutural no processo comunicativo, ou seja, a função comunicativa do elemento” (Martelotta; Kenedy, 2003, p. 19).

Nesse aspecto, a Linguística desenvolvida pelos linguistas de Praga previa a natureza funcional do sistema linguístico, que se prestava a desempenhar as funções pretendidas em determinada atividade comunicativa. Assim, a depender da ênfase dada à *função*, o estruturalismo pode se constituir tanto *formalista* quanto *funcionalista*. No polo formalista “a análise ressalta a forma linguística, ficando sua função num plano secundário”. Já o polo funcionalista define que “a função que a forma linguística desempenha no ato comunicativo tem papel predominante” (Martelotta; Kenedy, 2003, p. 19).

Os estudos formalistas orientam linguistas da Escola de Copenhague e estudiosos como Rask, Madvig, Noreen, Jespersen, Hjelmslev, entre outros. Além disso, essa base dos estudos linguísticos influenciou o descritivismo americano de Bloomfield, Trager, Bloch, Harris, Fries e as abordagens gerativistas de Noam Chomsky (Martelotta; Kenedy, 2003, p. 20). Assim, é possível afirmar que, por priorizar a análise da forma linguística, o polo formalista volta-se para a concepção de língua como entidade autônoma e suficiente por si mesma, em que as análises classificatórias da estrutura da forma valem por si, sem considerar contexto de comunicação e propósitos comunicativos. Tal concepção pressupõe que o estudo da língua deve se satisfazer com a análise da estrutura da língua como um conjunto de regras que se unem mecanicamente e que possuem sentido imanentemente linguístico.

A visão funcionalista, por sua vez, tem na *Escola Linguística de Praga* seus mais representativos estudos e desenvolvimento, sendo seus maiores representantes Troubetzkoy, o pensador russo Roman Jakobson e o linguista praguense Wilhem Mathesius, além da aceitação de outros linguistas como Halliday e Dik (Neves, 1997, p. 17). Ao compreender que a natureza funcional da linguagem exige que o seu foco de estudo esteja pautado nos usos reais da língua, o polo funcionalista dirige suas atenções para a análise e reflexão da língua em funcionamento, pois entende que o sistema linguístico por si é insuficiente para produzir sentido. Desse modo,

---

<sup>29</sup> Karl Bühler (1879-1963), linguista, com filósofo, psicólogo e psiquiatra alemão, um dos representantes do pensamento da linguística moderna. Determina o sistema tripartido de funções da linguagem, em que estabelece as bases para o amadurecimento de uma teoria da comunicação concebida posteriormente por um de seus discípulos, o linguista russo Roman Jakobson. Dentre suas obras destaca-se *Sprachtheorie* (Teoria da Linguagem (1934)).

<sup>30</sup> Neves (1997) destaca as concepções do termo função na visão de autores como Martinet (1994), Danes (1987), Dillinger (1991), Anscombe e Zaccaria (1990), Halliday (1973), entre outros. A autora observa que esse termo corresponde a diferentes conceitos atribuídos por diferentes linguistas funcionalistas, no entanto, sempre correlacionados ao propósito comunicativo e ao contexto discursivo.

é passa-se a considerar tanto o aspecto formal quanto o funcional da linguagem, entendendo as funções que as formas linguísticas desempenham em determinado discurso.

Sendo assim, ao se preconizar que a investigação sobre a linguagem deve priorizar a análise das funções desempenhadas pelas formas linguísticas, o funcionalismo linguístico elabora uma perspectiva social da própria linguagem que se estrutura e se organiza associada às relações que permeiam a atividade de uso e comunicação pela linguagem, sendo uma “corrente linguística que, em oposição ao estruturalismo e gerativismo, se preocupa em estudar a relação entre a estrutura gramatical das línguas e os diferentes contextos comunicativos em que elas são usadas” (Cunha, 2008, p. 157).

**Quadro 6** – Concepções de língua

<b>ESTRUTURALISTA/FORMALISTA</b>	Língua homogênea; Estudo da forma; Língua como abstração; Estudo da estrutura (independência da estrutura); Sujeito ideal (SAUSSURE, BLOOMFIELD, CHOMSKY); Autonomia da Estrutura.
<b>FUNCIONALISTA</b>	Língua heterogênea; Estudo da língua em uso; Relação entre forma e contexto; Relação entre os sujeitos (FUNCIONALISMO, SOCIOLINGUÍSTICA, BAKHTIN).

Fonte: Elaborada pela autora

Nesse sentido, passa-se a considerar o contexto em que a atividade verbal ocorre, os indivíduos que se apropriam do sistema, as escolhas que fazem e o seu propósito comunicativo. Marcamos assim, um avanço em relação as abordagens formalistas – estruturalismo e gerativismo –, uma vez que passamos a conceber a linguagem como um instrumento de interação social, buscando no contexto discursivo, em situações reais, a motivação para os fatos da língua, encontrando caminhos que passam a possibilitar que a mudança foi compreendida em um quadro de diversidade de usos linguísticos, como propões atualmente a Sociolinguística Variacionista/ Sociolinguística Quantitativa, a ser discutida na *seção 3.2*.

### 3.2 A Sociolinguística e a sistematicidade da mudança

Até então vimos que a Linguística, ao se constituir como ciência autônoma, a partir das reflexões de Ferdinand de Saussure (2006 [1916]), tem como objeto de estudo a *langue*, estrutura homogênea. Essa concepção de língua é repensada pelos funcionalistas que passam a compreendê-la a partir funções comunicativas – ou seja, a função passa a determinar a forma, divergindo da concepção formalista (estruturalista e gerativistas) em que temos a organização da língua motivada fatores inerentes a própria estrutura.

Portanto, será a partir de meados do século XX que a área da Linguística sofrerá mudanças significativas, em que estudos linguísticos passam a se interessar não pelo sistema da língua em si, mas também pelo seu uso. Dentre esse campo de investigação destacamos nesta seção a Sociolinguística que, em oposição ao formalismo, se preocupa em estudar a relação entre língua e sociedade, debruçando-se sobre investigações da variação e da mudança linguísticas, compreendendo-as como inerentes ao sistema linguístico, cujo funcionamento é motivado por tanto por fatores de ordem interna à língua quanto por fatores externos.

Antecipadamente destacamos que este modelo teórico fundamentará a constituição do *corpus* do Projeto Estudos Linguísticos e Históricos do Sertão (ELiHS) bem como o recorte da amostra estudada/analísada, comunidade afro-brasileira de Volta do Angico, e as possíveis interpretações no que tange ao fenômeno de variação na concordância nominal de número. Dito isso, importante retomarmos algumas questões antes de adentrarmos nas concepções teórico-metodológicas da Sociolinguística.

Calvet (2002, p. 13-14), no seu livro intitulado *Sociolinguística: uma introdução crítica*, apresenta criticamente os progressos da Linguística e inicia afirmando que foi o linguista francês Antoine Meillet (1866-19936), contemporâneo de Saussure, quem insistiu em numerosos textos, no caráter social da língua, definindo-a como um fato social. Porém, com a publicação póstuma do *Curso de linguística geral*, Meillet afirma, em resenha que faz do livro, que ao separar a variação linguística das condições externas de que ela depende, Ferdinand de Saussure a priva da realidade; ele a reduz a uma abstração que é necessariamente inexplicável. Calvet (2002, p. 15) aponta que

A afirmação do caráter social da língua que se verifica em toda a obra de Meillet implica ao mesmo tempo a convergência de uma abordagem *interna* e de uma abordagem *externa* dos fatos da língua e de uma abordagem *sincrônica* e *diacrônica* desses mesmos fatos. Quando Saussure opõe linguística interna e linguística externa, Meillet as associa; quando Saussure distingue abordagem sincrônica de abordagem diacrônica, Meillet busca explicar a estrutura pela história [...]. Enquanto Saussure

busca elaborar um modelo abstrato da língua, Meillet se vê em conflito entre o *fato social* e o *sistema que tudo contém* [...].

O autor (2002, p. 32) ainda assinala que Meillet já afirmava que é preciso buscar a explicação da irregularidade das variações linguísticas nas flutuações da composição social da comunidade linguística e isso é retomado por Labov ao publicar, em 1966, seu estudo sobre a estratificação social do /r/ nas grandes lojas de Nova York. Porém, Calvet (2002, p. 33) ressalva que

Meillet, comparatista de alto nível, trabalhou sobretudo com línguas mortas, enquanto Labov trabalha continuamente com situações contemporâneas concretas, enfrenta problemas de metodologia da pesquisa, em suma, constrói um instrumento de descrição que tenta ultrapassar, integrando-os, os métodos heurísticos da linguística estrutural.

Nessa perspectiva, a Sociolinguística, proposta inicialmente por William Labov, propõe uma reação à ausência do componente social no modelo gerativo e veementemente insiste na relação entre língua e sociedade e na possibilidade de se sistematizar a variação existente na própria língua falada, fixando um modelo de descrição e interpretação do fenômeno linguístico, conhecido como Sociolinguística Variacionista ou Teoria da Variação, de grande impacto na linguística contemporânea.

Dentre os trabalhos publicados pelo autor, destaca-se o publicado em 1963<sup>31</sup>, intitulado *The social motivation of a sound change*, um dos célebres trabalhos de Labov, em que trata sobre o inglês falado na ilha de Martha's Vineyard, no estado de Massachusetts– Estados Unidos, destacando o papel decisivo dos fatores sociais na explicação da variação linguística, isto é, da diversidade linguística observada. Em 1966, Labov realiza um estudo intitulado *The social stratification of english in New York City* sobre a estratificação social do inglês em Nova York, a partir do qual fixa um modelo de descrição e interpretação do fenômeno linguístico no contexto social de comunidades urbanas – conhecido como Sociolinguística Variacionista ou Teoria da Variação, de grande impacto na linguística contemporânea. Neste contexto, surgiu a necessidade e a importância de se definir, na teoria da variação, a noção de vernáculo, permitindo repensar alguns procedimentos metodológicos.

Com base em estudos empíricos, os linguistas Weinreich, Labov e Herzog lançaram o livro *Fundamentos Empíricos para uma Teoria da Mudança Linguística* (1968), em que se propõem estudos sobre a gramática de uma comunidade de fala, buscando explicar o

---

<sup>31</sup> Primeiro estudo quantitativo realizado por William Labov, em que relacionou fatores como idade, sexo, ocupação, origem étnica e atitude em relação ao comportamento linguístico dos nativos da ilha Martha's Vineyard, no que se referia à pronúncia de determinados fones do inglês (as vogais dos ditongos [ay] e [au]), ao qual se seguiram estudos sobre a estratificação social do inglês falado na cidade de Nova York.

funcionamento dessa gramática a partir de seus aspectos internos e externos. Hoje esse texto tornou-se um importante material, uma vez que nele há trabalhos empíricos de dialectologia e sociolinguística com implicações para a compreensão da mudança, delineando uma estratégia geral que se apoia justamente nesses fundamentos empíricos para estudar a mudança linguística.

Através deste livro, de estudos de bases empíricas, foi possível compreender melhor a heterogeneidade linguística do português do Brasil, instaurando uma nova perspectiva de linguagem, uma vez que ao enfatizar o lugar central da variação e da mudança nas línguas naturais e explicitar os diversos aspectos envolvidos nesses processos, os autores abrem espaço, como afirmam as autoras Duarte e Paiva no posfácio do livro, para o desenvolvimento de uma modelização orientada por questões precisas.

Diante do legado deixado pelos autores podemos entender que a sociolinguística é um espaço de investigação interdisciplinar que estuda e correlaciona aspectos dos sistemas linguísticos e dos sistemas sociais, através dos empregos concretos da língua, dentro das comunidades de fala. Nesse sentido, a concepção e o alcance do modelo sociolinguístico passam ser, ao mesmo tempo, sincrônicos e diacrônicos, já que toda mudança implica uma variação. Nesse novo modelo, que surge na década de 1960, a estrutura de uma língua somente será totalmente entendida à medida que se compreendam efetivamente os processos históricos de sua configuração.

Nesse contexto, Weinreich, Labov, Herzog, em sua introdução afirmam que *Empirical Foundations for a Theory of Language Change* (2006 [1968])

[...] se baseia na observação de que as teorias estruturais da língua, tão frutíferas na investigação sincrônica, sobrecarregaram a linguística histórica com um fardo de paradoxos que não foram totalmente superados. Ferdinand Saussure, ao assentar as fundações do estudo sincrônico, estava ciente da correspondente intratabilidade da mudança linguística, e claramente resignado com isso. Mas entre a maioria dos linguistas depois de Saussure, a escolha entre estudar ou a estrutura ou a história das línguas não ficou bem assentada (Weinreich; Labov; Herzog, 1968, traduzido por Bagno, 2008, p. 33).

No que tange a aproximação entre a Sociolinguística e a Linguística Histórica, Oliveira (2006) comenta a existência de uma relação estreita entre essas áreas, uma vez que, estudando a variação da língua dentro da estrutura social da comunidade em que ocorre, identifica e prevê processos de mudança em curso, que, obviamente, são decorrentes dessas variações. A autora ainda afirma que “sendo a língua um meio de interação entre o indivíduo e a sociedade em que ele atua, está claro que há forças sociais que agem sobre essa interação” (Oliveira, 2006, p. 45), sendo assim “no modelo da sociolinguística [...] o que se oferece ao falante não é um sistema homogêneo, unitário e imutável, que se impõe de forma irredutível, mas um sistema

heterogêneo sobre o qual o falante atua de acordo com as disposições estruturadas em que a prática lingüística se atualiza” (Lucchesi, 2004, p. 172).

Dessa forma, os estudos que hoje buscam o relacionamento casual entre língua e sociedade, são frutos da proposta de análise contida em *Empirical Foundations for a Theory of Language Change* (2006 [1968]), em que se propõem estudos sobre a gramática de uma comunidade de fala. Hoje esse texto tornou-se um importante material, uma vez que

Nele – depois de uma revisão crítica do pensamento neogramático e do estruturalismo saussuriano, em que apontam os paradoxos introduzidos pelos imanentistas na reflexão histórica com sua concepção homogeneizante de língua – aqueles autores, revendo trabalhos empíricos de dialectologia e sociolinguística com implicações para a compreensão da mudança, delineiam uma estratégia geral que se apoia justamente nesses fundamentos empíricos para estudar a mudança linguística (Faraco, 2005, p. 189).

Faraco (2008) – na apresentação da versão brasileira dos *Fundamentos Empíricos* – afirma que os textos de Weinreich, Labov, Herzog apresentam complexidade, ao trazer desafios para os estudos da mudança linguística e ao atentar para o fato de que os fatores que desencadeiam e restringem as mudanças são certamente mais diversificados em comunidades linguísticas contemporâneas do que em sociedades pré-urbanas, que até então era foco de parte das investigações históricas tradicionais em linguística. Sendo assim, a Sociolinguística passa a ser compreendida como um modelo que

[...] acomoda os fatos do uso variável e seus determinantes sociais e estilísticos não só leva a descrições mais adequadas da competência lingüística, mas também suscita naturalmente uma teoria da mudança lingüística que ultrapassa os estereis paradoxos contra os quais a lingüística histórica vem lutando há mais de meio século. (Weinreich, Labov, Herzog, 2006 [1968], p. 34).

Dessa forma, a heterogeneidade linguística sai do plano hipotético dos estudos saussureanos e ganham espaços nas investigações sociolinguísticas, que passa a entender a língua dentro do seu contexto de variabilidade e mudança. Nesse novo contexto teórico, cinco são os problemas apontados por Weinreich, Labov, Herzog (2006 [1968], p. 121-125) quando se trata de qualquer teoria da mudança linguística. São eles:

- (i) **O problema dos fatores condicionantes (*the constraint problem*)** – A teoria precisa determinar o conjunto de mudanças possíveis e as condições possíveis para a mudança por meio de um estudo minucioso de mudanças em progresso;
- (ii) **O problema da transição (*the transition problem*)** – As mudanças têm uma distribuição contínua através de sucessivas faixas etárias da população;



- (iii) **O problema do encaixamento (*the embedding problem*)** – As mudanças linguísticas devem ser vistas como encaixadas no sistema linguístico como um todo. Há dois tipos de encaixamento:
- *Encaixamento na estrutura linguística* – Apresenta um modelo de língua em que se tem: a) estratos discretos, coexistentes, definidos pela co-ocorrência estrita, que são funcionalmente diferenciados e conjuntamente disponíveis a uma comunidade de fala; b) variáveis intrínsecas, definidas por co-variação com elementos linguísticos e extralinguísticos, o que permitiu afirmar que um conjunto limitado de variáveis num sistema altera seus valores modais gradualmente de um pólo a outro. Portanto, as variantes das variáveis podem ser contínuas ou discretas;
  - *Encaixamento na estrutura social* – A estrutura linguística está encaixada no contexto mais amplo da comunidade de fala. Portanto, variações sociais e geográficas são elementos intrínsecos da estrutura, uma vez que os fatores sociais pesam sobre o sistema como um todo;
- (iv) **O problema da avaliação (*the evaluation problem*)** - O linguista deve estabelecer empiricamente os correlatos subjetivos da avaliação dos diversos estratos e variáveis numa estrutura heterogênea. Além disso, o nível de consciência social é uma propriedade importante da mudança linguística que tem de ser determinada diretamente;
- (v) **O problema da implementação (*the actuation problem*)** - O processo global da mudança linguística pode envolver estímulos e restrições tanto da sociedade quanto da estrutura da língua. O problema em torno da implementação advém do grande número de fatores que influenciam a mudança;

Ainda cabe salientar que a Teoria da Variação considera a língua em seu contexto sociocultural, uma vez que parte da explicação que a heterogeneidade emerge nos usos linguísticos concretos, reflexo de fatores externos ao sistema linguístico e não só nos fatores internos à língua. Portanto, como observou Mollica (2004, p. 10), “ela parte do pressuposto de que toda variação é motivada, isto é, controlada por fatores de maneira tal que a heterogeneidade se delinea sistemática e previsível”.

Desse modo, um estudo sociolinguístico visa à descrição fundamentada de um fenômeno variável, tendo como objetivo analisar, descrever e sistematizar variantes linguísticas usadas por uma mesma comunidade de fala. Para tanto, calcula-se a influência que cada fator, interno ou externo ao sistema linguístico, possui na realização de uma ou de outra variante. Ao formalizar esse cenário, a análise sociolinguística busca estabelecer a relação entre o processo de variação que se observa na língua em um determinado momento (isto é, *sincronicamente*) com os processos de mudança que estão acontecendo na estrutura da língua ao longo do tempo (isto é, *diacronicamente*).

Nesse sentido, a concepção e o alcance do modelo sociolinguístico são, ao mesmo tempo, *sincrônicos* e *diacrônicos*, já que toda mudança implica uma variação. Nesse novo modelo, que surge na década de 1960, a estrutura de uma língua somente será totalmente entendida à medida que se compreendam efetivamente os processos históricos de sua configuração, já que se torna salutar os estudos que tratam da relação de aspectos de natureza sócio-históricas com aspectos de natureza linguística, principalmente no campo da *variação e mudança linguística no Português do Brasil*.

Neste contexto, não podemos deixar de frisar a importância da proposta desta tese, que busca conciliar aspectos de natureza sócio-histórica com aspectos de natureza linguística, por meio de uma análise variacionista laboviana da concordância nominal de número no português falado em comunidades de afro-brasileiras do Território de Identidade de Irecê, apresentando resultados que endossam as pesquisas científicas, numa proposta de interiorização linguística que favorece a discussão sobre a formação da realidade sociolinguística brasileira, colaborando para o desafio de Weinreich, Labov e Herzog (2006 [1968]) de buscar um modelo formal capaz de acomodar a heterogeneidade ordenada como uma realidade inerente às línguas e não como um fundamento marginal.

Não poderíamos deixar de expor, mesmo que ainda de forma superficial, que ao compreendermos os processos de variação perpassam pela perspectiva social, a Sociolinguística, segundo Eckert (2012), pode ser melhor compreendida em três ondas, que refletem perspectivas distintas de análise. A autora centra suas discussões na considerada terceira onda da Sociolinguística, em que preocupações sociais, parecem ser mais evidentes, na direção de uma mudança social, perspectiva que não será assumida nesta tese, mas abre caminhos para possíveis e futuros desdobramentos quando pensamos em comunidades de prática de comunidades rurais afro-brasileiras, autointituladas enquanto remanescentes de quilombos.

Neste contexto de categorizar em ‘ondas’ o tratamento teórico e metodológico em relação aos processos de variação, Eckert (2012) compreende que na Sociolinguística Variacionista, ou de primeira onda, a variação estilística é o resultado do pertencimento do falante a categorias macros sociais e ao monitoramento da fala em contextos estilísticos específicos, tendo o falante como um agente passivo na construção de seu comportamento estilístico<sup>32</sup>. Segundo a autora, a análise da mudança linguística em espaços sociais, configura o que se considera a primeira onda dos estudos da Sociolinguística, tendo os estudos da variação duas ondas subsequentes.

A autora descreve a primeira onda de estudo de variação, caracterizada através de estudos desenvolvidos por Labov (1966; 1994; 2008 [1972]), demonstrando o processo de estratificação social utilizado em suas pesquisas, as quais foram realizadas através de entrevistas gravadas, representativas da língua falada, constituindo amostras de fala. Neste contexto, a autora afirma surgir a necessidade e a importância de se definir, na teoria da variação, a noção de vernáculo, permitindo repensar alguns procedimentos metodológicos. Sendo assim, Eckert (2012), através das pesquisas elencadas/descritas, afirma que os estudos de primeira onda interpretaram a significado social da variação, com base em uma compreensão geral das categorias que serviu para selecionar e classificar informantes e não através do conhecimento direto desses falantes e suas comunidades.

Na tentativa de atender a essas lacunas deixadas pela primeira onda dos estudos de variação, surge a segunda onda, que passou a considerar métodos etnográficos, no intuito para se aproximar mais dos falantes e suas comunidades de fala. Assim, busca-se uma relação direta entre a dinâmica social que deu origem a essas categorias sociais e uso de variáveis linguísticas, em que temos uso do vernáculo como expressão local ou de classe de identidade. Nesse sentido, Eckert (2012) afirma que Milroy (1980), inspirado no trabalho de Gumperz (1982), que inaugura a segunda onda com um estudo de variação fonológica em redes sociais em Belfast, procurando forças positivas no uso vernacular da classe trabalhadora de Belfast, o que permitiu levantar a hipótese de que classe trabalhadora teria um forte poder local de imposição de normas, o que permitiria correlacionar os tipos de rede dos indivíduos com o uso de variáveis vernaculares.

---

<sup>32</sup> O estudo quantitativo realizado por William Labov (1963) em Martha's Vineyard marca a primeira onda da Sociolinguística, em que temos a análise da pronúncia de / ay / e resultados apontavam que os falantes da ilha vinham acompanhando tendência continental para baixar o núcleo para [a], o que gerava uma característica estilística marcante aos falantes da ilha, configurando uma identidade peculiar

Nesse sentido, a autora apresenta outros estudos nessa perspectiva – Cheshire (1982); Rickford (1986); Holmquist (1985), tratando também do fato aparente de os adolescentes liderarem mudança de som e no uso do vernáculo, o que levanta a questão do papel da classe nos adolescentes, o que leva Eckert (1989, 2000) a realizar um estudo etnográfico de adolescentes nas escolas secundárias, predominantemente brancas, suburbana de Detroit. Esses estudos indicam que correlações mais amplas de classe não são simplesmente as consequências da educação, ocupação e renda, mas refletem as dinâmicas enraizadas em práticas e ideologias e essas, por sua vez, são moldadas por classe.

Esses fatos refletem que os estudos da linguagem caracterizam perspectivas teóricas inteiramente novas, que culminarão na guinada da terceira onda da Sociolinguística, visto que enquanto as duas primeiras ondas viram o significado da variação como precipitação acidental de espaço social, a terceira onda o vê como uma característica essencial da linguagem, compreendendo que a medida que a sociedade sofre mudanças sociais, essas são perpassadas por seus falantes, não cabendo as teorias linguísticas engessarem suas categorias de análises (Eckert, 2012).

Para esclarecer sobre a perspectiva de análise da terceira onda, Eckert (2012), mais uma vez, cita pesquisas que se desdobram nesta nova vertente, focando, neste momento, em descrições estilísticas de comunidades de práticas. Assim, chega à conclusão que na mudança da primeira para a terceira onda nos estudos de variação, temos uma nova visão, no que tange a relação entre linguagem e sociedade, passando, na terceira onda, a uma ênfase em práticas estilísticas, perpassadas por ideologias, compreendendo-as como produção de diferenciação em contexto de variação linguística, o que culmina em uma teoria linguística ainda mais social.

Assim, em cada onda sociolinguística teremos um olhar e um aparato teórico metodológico que contemple os objetivos centrais de pesquisa realizados no campo linguística. Nesta tese, centraremos nossas análises, em relação aos processos de variação na concordância nominal de plural, na Sociolinguística Variacionista, ou de primeira onda (Eckert, 2012), compreendendo que os processos de variação são condicionados por fatores extralinguísticos, utilizando-nos dos métodos de coleta de dados quantitativo, que serão melhores descritos na seção que segue.

### 3.2.1 A Sociolinguística Quantitativa

A tradicional pesquisa sociolinguística, ou de primeira onda (*cf.* seção 3.2), circunscreve-se na relação entre língua e sociedade e na possibilidade de investigar e descrever

sistematicamente a variação existente em contextos específicos de fala, no uso real da língua. Essa base teórica deve-se ao americano William Labov, a quem se deve as primeiras investigações com base nos postulados da sociolinguística, que também é conhecida como “sociolinguística quantitativa”, pois em se tratando dos processos de variação e mudança linguísticas, torna-se crucial se operar com números e tratamento estatístico dos dados coletados. No que tange a análise quantitativa, Guy e Zilles (2007) dedicam-se intensamente à tarefa de oferecer ao leitor subsídios para uma análise quantitativa adequada aos princípios teórico-metodológicos da sociolinguística quantitativa variacionista.

Para o desenvolvimento de uma análise sociolinguística, em que seja possível abarcar fatores linguísticos e extralinguísticos, é necessário dados quantitativos para que seus resultados sejam representativos e possam, realmente, indicar os caminhos da variação e/ou mudança de certas variantes linguísticas. Para tanto, se faz necessário todo um cuidado metodológico já abordado por Tarallo (2007 [1997]), Guy e Zilles (2007), Mollica e Braga (2008) e outros pesquisadores, que vai desde a coleta de dados a interpretação dos resultados. Como já afirma Guy (2007, p. 19):

Para desvelar tanto a estrutura linguística quanto a estrutura social, devemos, necessariamente, coletar grande quantidade de dados de muitos indivíduos; conseqüentemente, devemos enfrentar problemas ligados a controle de qualidade e confiabilidade, a manuseio e apresentação de dados, e a interpretação e inferência.

O autor segue comentando que parece sensato afirmar que toda pesquisa dialetal, seja ela de ordem geográfica ou social é essencialmente quantitativa, uma vez que estudos dialetais anteriores podem comprovar que dificilmente se fala em termos categóricos sobre as propriedades dos dialetos. O autor ainda explica que isso acontece diante da heterogeneidade da língua falada, já que as isoglossas não formam grupos perfeitos, o que implica dizer que todo dialeto está propício a variação linguística (Guy, 2007, p. 20). Portanto, cabe a sociolinguística fazer a abordagem linguística e social dos mais diversos dialetos, tomando como suporte metodológico a análise quantitativa dos dados, podendo assim controlar o grau de variabilidade do fenômeno em análise, e apresentá-los através dos mais variados recursos gráficos (tabelas, gráficos, fluxogramas).

Numa análise quantitativa, o pesquisador busca sempre os números, as probabilidades, as estatísticas, para que seja possível apreender a sistematicidade da variação linguística, em análise, uma vez que esses resultados refletirão o uso da língua em um contexto social e linguisticamente heterogêneo. Dessa forma, entendemos que na abordagem da sociolinguística quantitativa

os padrões linguísticos são descritos, e explicados (na medida em que objetos dessa natureza podem ser explicados), em termos de uma gramática de regras variáveis que operam com probabilidades associadas a fatores sociais e restrições hierarquizadas da estrutura linguística (Lucchesi, 2004, p. 196)

De qualquer modo, Guy (2007) define três fases numa análise quantitativa. São elas:

- (i) **Coleta de dados** – momento em que se organiza uma amostra de dados, a qual deve apresentar altos índices de confiabilidade e representatividade, para que os dados não sejam tendenciosos;
- (ii) **Redução e apresentação de dados**– abordagem que busca formas que minimize detalhes sem importância e apresente o que realmente é relevante à análise, sem que isso torne os dados tendenciosos ou distorcidos, ou até mesmo para não incorrer no risco de obscurecer/omitir dados/fatos importantes;
- (iii) **Interpretação e explicação de dados**– representa o objetivo final de qualquer estudo quantitativo dialetal, uma vez que os números/estatísticas apenas direcionam as análises e permitem interpretações e explicações dos fenômenos linguísticos.

Todo o percurso, da coleta a interpretação –que serão melhores descritos na *seção 4.2.2* –, ganhou um grande auxílio das ciências exatas e computacionais, que na tentativa de facilitar o trabalho de quantificação de dados elaboraram programas computacionais que ofereceram um tratamento estatístico aos dados sociolinguísticos, transformando todo o amontoado de ocorrências em probabilidades estatísticas. A inserção de programas estatísticos desenhados, especificamente, para lidar com dados de variação linguística trouxe ao campo da Sociolinguística revestiram as análises de uma maior precisão e confiabilidade.

Todos esses recursos oferecidos pelo mundo moderno têm vindo a somar na realização pesquisas mais confiáveis e pensarmos os processos de variação linguística de forma mais sistematizada, a partir de análise multivariadas. Em meios aos avanços e modernização da sociolinguística quantitativa, podemos afirmar que

antes do advento da metodologia de quantificação, a variação linguística era considerada secundária, aleatória ou mesmo impossível de ser cientificamente apreendida. O uso de métodos estatísticos, contudo, tem permitido demonstrar o quanto central a variação pode ser para o entendimento de questões como identidade, solidariedade ao grupo local, comunidade de fala, prestígio e estigma, entre tantas outras (Guy; Zilles, 2007, p. 73).

No que tange aos programas computacionais mais tradicionais temos o *Variable Rule Analysis*<sup>33</sup> (VARBRUL), que é um tipo de análise multivariada amplamente empregada em estudos de variação linguística. Este programa busca separar, quantificar e testar a significância dos efeitos dos fatores (sociais ou linguísticos) em uma variável linguística (Guy, 2007, p. 33). Desde a criação do VARBRUL, muitas foram as alterações realizadas no programa para que esse se torne, a cada dia, mais prático, confiável e com mais ferramentas operacionais, o que tem colaborado na aplicação de testes, cruzamentos de dados.

O pacote computacional *Varbrul* foi atualizado para *Goldvarb*, que já passou por algumas atualizações, sendo o *Goldvarb X*<sup>34</sup> a versão utilizada no tratamento estatísticos dos dados desta tese. Mais recentemente foi elaborado por Daniel Ezra Johnson (Johnson, 2009) um programa de modelagem estatística de regressão logística com resposta binária, a linguagem de programação denominada R (R CORE TEAM, 2020), que por meio de suas interfaces reporta quantitativamente os condicionadores da escolha das variantes em análise – o programa será descrito na *seção 5.2.2*.

No desenvolver deste capítulo vimos que vários são os caminhos teóricos no desenvolvimento de investigações linguísticas (*cf.* seção 3.1), esses caminhos condicionam outras tantas possibilidades metodológicas, extremamente importantes na pesquisa científica. No encontro entre teoria e métodos somos direcionados a novas escolhas, a dos recursos computacionais que favoreçam a concretização efetiva dos objetivos fulcrais da pesquisa. Neste campo de escolhas e enquadramentos metodológicos optamos pelo R (R CORE TEAM, 2020), por meio de uma de suas interfaces, o *R Studio* (EQUIPE RSTUDIO, 2020), entendendo que este pacote computacional atende ao cumprimento de todas as etapas da pesquisa de cunho sociolinguísticos, e para a análise empírica que propomos realizar atende as variáveis elencadas como possíveis condicionantes do processo de variação do fenômeno em análise (*cf.* seção 5.3).

---

<sup>33</sup> Análises de regras variáveis.

<sup>34</sup> O *Goldvarb X* é um software gratuito, desenvolvido por Sankoff, Tagliamonte e Smith (2005), para análise de *corpus*, executável em versões para o *Windows*, *Linux* e *Macintocsh*. Seu arquivo tem aproximadamente 2,66MB, por isso é considerado leve e dispensa a necessidade de instalação, o que torna a utilização mais fácil mesmo para usuários iniciantes. O programa está disponível na página <http://individual.utoronto.ca/tagliamonte/goldvarb.html>.

**PARTE II – CONTRIBUIÇÕES PARA PROPOSTA DE INTERIORIZAÇÃO  
LINGUÍSTICA**



#### 4 EM DIREÇÃO AO SERTÃO E DO PRESENTE PARA O PASSADO: ELiHS, UM PROJETO, UM *CORPUS* E UMA PLATAFORMA *ONLINE*

*“Da minha aldeia vejo quanto da terra se pode ver do Universo...  
Por isso a minha aldeia é tão grande como outra terra qualquer  
Porque eu sou do tamanho do que vejo  
E não do tamanho da minha altura...”*

*(Pessoa, 1993 [1946])*

Neste capítulo, descreveremos particularidades do Projeto intitulado *Estudos Linguísticos e Históricos do Sertão* (ELiHS), o qual foi idealizado e executado no âmbito desta tese, bem como suas possíveis contribuições para a história do português popular no Sertão da Bahia, retomando alguns aspectos sócio-históricos da região territorial de Irecê – BA. Não deixando de reconhecer outros Projetos pioneiros que têm colaborado no processo de salvaguardar este patrimônio imaterial, o português popular do Brasil, investiga, em perspectiva sincrônica, do presente para o passado, seu funcionamento.

O Projeto ELiHS norteia-se na concepção de que o contato linguístico e cultural esteve presente na socio-história brasileira e de que o Português Brasileiro nasce dessa diversidade (Houaiss, 1985, p. 91), perspectiva já discutida e fundamentada no *Capítulo 2* desta tese. O que ainda se faz importante frisar, e o Projeto ELiHS toma como objetivo balizador, é a compreensão dos afro-brasileiros como um dos principais difusores do português vernacular brasileiro (Mattos e Silva, 2004a). Apoiando-nos nessa perspectiva, neste capítulo, daremos visibilidade às comunidades rurais afro-brasileiras que estão “para além do litoral” e quase sempre estão à margem das pesquisas científicas, mas que **existem** e **resistem** no Território brasileiro, em especial o de Irecê, que agrega oficialmente um dos maiores quantitativos de comunidades, intituladas, pela Fundação Cultural Palmares, de Remanescentes de Quilombos (*cf.* site da Fundação Palmares<sup>35</sup>), totalizando 148 certificações até o primeiro semestre do ano de 2023. Esse quantitativo, enfaticamente, é um convite à pesquisa científica e, portanto, um desafio para o Projeto ELiHS, quanto a tornar essas comunidades afro-brasileiras, em uma analogia a nossa epígrafe introdutória, tão grandes como outra comunidade qualquer, fazendo-as presentes no mapa sociolinguístico brasileiro, marcando seu lugar e suas vozes na formação da nossa língua. Que o Projeto ELiHS, um dos produtos desta tese, seja mais um importante passo dado ao processo de interiorização linguística, e que o Território de Irecê, a partir da composição de

---

<sup>35</sup> Acesso através do link: <https://www.gov.br/palmares/pt-br>

*corpora* orais, seja contemplado com pesquisas linguísticas descritivas, nas diversas perspectivas teóricas, que colaborem na compreensão da formação do português brasileiro, em especial em sua vertente popular.

No tocante às comunidades tradicionais afro-brasileiras, autodeclaradas enquanto remanescentes de antigos quilombos, é facilmente notável que o inexistir de informações assola nosso cenário atual. Pouco ou nada temos registrado sobre essas comunidades, o que torna urgente a tentativa de recuperação e resgate dos fragmentos de informações que possam permitir a compreensão do que está acontecendo no processo de distribuição dos remanescentes de antigos quilombos do Brasil (Anjos, 1999), o que diretamente reverbera no conhecimento sobre os aspectos sócio-históricos fundamentais para compreensão da formação do português brasileiro (Mattos e Silva, 2004a). Na tentativa de colaborar no preenchimento dessas lacunas seculares, aqui também apresentaremos dados informacionais inéditos que mapearão e caracterizarão, de forma ainda generalizada, as comunidades rurais afro-brasileiras situadas no Território de Irecê. O inexistir de informações sobre essas comunidades dificulta o acesso e a visibilidade frente aos poderes públicos; assim, o Projeto ELiHS, ainda que timidamente, busca, através da sua plataforma *online*, gratuita, direcionar os holofotes, não só científicos, como sócio-históricos, para formação e permanência das 148 comunidades que lutam diariamente por uma identidade marcada por sua ancestralidade.

Iniciar o mapeamento das comunidades afro-brasileiras do Território de Irecê traz à tese um impacto interventivo que aproxima de forma direta a comunidade externa da academia, visto que o mapeamento e registro de seus territórios potencializa a luta pela garantia de direitos e supre lacunas existentes nos dados oficiais. No entanto “não podemos perder de vista que um mapa não é o território, mas que nos produtos da cartografia estão as melhores possibilidades de representação e leitura da história do território” (Anjos, 2011, p. 17).

Ao contemplarmos os aspectos sócio-históricos e de mapeamento, discutiremos os critérios metodológicos adotados na constituição das amostras sincrônicas do banco de dados do Projeto ELiHS, na tentativa de apresentarmos/disponibilizarmos amostras confiáveis, a partir de critérios metodológicos bem delimitados, comungando com a ideia de Freitag (2017), e outras importantes referências, de que a abordagem metodológica perpassa pelo prisma da pesquisa Sociolinguística.

Por fim, apresentaremos a estrutura do *site* do Projeto, uma plataforma *online* inteiramente gratuita, através da qual armazenaremos e disponibilizaremos: fotografias representativas das comunidades; dados sócio-históricos inéditos; gravações de áudio em

formato *MP3*; transcrições em versão *DOC* e *XML (Extensible Markup Language)*. Uma base de dados inédita e gratuita torna possível a ampliação do conhecimento da realidade sociolinguística do Interior da Bahia e impacta diretamente o desenvolvimento de projetos de pesquisa e extensão, em especial, do curso de Letras, Língua Portuguesa e Literaturas da Universidade do Estado da Bahia, *campus XVI - Irecê*, instituição sede do Projeto ELiHS, bem como do curso de Letras Vernáculas da Universidade Estadual de Feira Santana (UEFS), do qual faz parte o principal projeto parceiro do ELiHS, o Projeto CE-DOHS.

Importante frisar que a proposta de criação de um novo banco eletrônico surge dentro de uma importante agenda de pesquisa, iniciada por outros importantes grupos de pesquisa, a citar o CE-DOHS, que sediará o *corpus* deste Projeto, bem como oferecerá o aparato tecnológico para o processo de edição em linguagem *XML* do *corpus*, através da ferramenta *eDictor* (Paixão de Souza; Kepler; Faria, 2009).

É nesse contexto de anseios e no afã de proporcionar material oral sincrônico para análise de descrições, especialmente no campo da linguística, que se delineia o Projeto *Estudos Linguísticos e Históricos do Sertão* (ELiHS), uma proposta pioneira no âmbito do Território de Irecê, que permanecerá como legado ao Departamento de Ciências Humanas e Tecnologias (DCHT) da Universidade do Estado da Bahia (UNEB), *campus XVI – Irecê/BA*. Que outras iniciativas como esta sujam na tentativa de salvuardarmos bancos de dados orais, patrimônio imaterial quando se trata da história do português popular brasileiro.

#### **4.1 Território de Identidade de Irecê e as comunidades rurais afro-brasileiras: uma breve caracterização sócio-histórica**

É sabido que a compreensão da realidade social e linguística do interior baiano, bem como de sua formação histórica, constitui um vasto campo de estudos ainda a ser, acertadamente, investigado por estudiosos das mais diversas perspectivas teóricas. No campo linguístico, Mattos e Silva (2008a) faz referência à importância de se extrair argumentos de dados de demografia histórica, de mobilidade sociocultural, geográfica, político-econômica, como fatores relevantes na recomposição de quadros sociolinguísticos de uma sociedade. Torna-se, então, indispensável ao pesquisador a avaliação dos dados que a história lhe legou, o que propomos para esta seção; afinal, diversas foram as formas de apropriação do sertão brasileiro. Sendo assim, as reflexões que se seguem são um convite para a construção de uma história social e linguística das comunidades rurais afro-brasileiras do Território de Identidade

de Irecê (doravante TII), “[...] obra que não pode deixar de ser coletiva e conjuntamente sonhada” (Mattos e Silva, 2004, p. 11).

Antes de adentrarmos nas questões propostas para esta seção, faz-se importante a explicação da posição assumida nesta tese de trabalharmos com a concepção de Território de Identidade, visto que o *locus* das amostras representativas do português rural afro-brasileiro que comporá o *corpus* do Projeto ELiHS situa-se nesse Território.

Segundo Santos<sup>36</sup> (2014, p. 13), nos últimos anos têm sido intensificadas as discussões voltadas para a promoção do desenvolvimento territorial no Brasil, em que se busca reforço em estudos sobre território sob o olhar das mais diversas ciências, na tentativa de superar a concepção de território apenas como espaço físico geográfico.

Na superação de entender o território apenas como espaço físico geográfico, outras correntes apresentam ideias colocando este (território) como espaço modificado pelo trabalho e que revela relações de poder, conforme defende Lefebvre (1978), Saquet (2010) e Raffestin (1993), que também o sinaliza como cena do poder e o lugar de todas as relações, ao argumentar que cada sociedade organiza seu espaço combinando elementos e, desta maneira, modela seu território, através da energia, da informação, de códigos, de estratégias, de objetivos e da comunicação, de forma relacional e multidimensional, por meio das relações sociais da vida de cada dia, ideias que passam a reconstruir o conceito de território ao longo dos tempos [...]. Assim, *supera-se também a compreensão de território como conceito político-administrativo, limitado às fronteiras e tendo o estado como única fonte de poder no controle espacial* (Santos, 2014, p. 14, grifo nosso).

Esse novo cenário passa a ser politicamente constituído a partir de 2003, com a criação do Programa Nacional de Desenvolvimento Sustentável dos Territórios Rurais (doravante PNDST), em que temos uma nova concepção de território adotada pelo Governo brasileiro. Cabe ressaltar que a política de desenvolvimento territorial, proposta no Brasil pela Secretaria de Desenvolvimento Territorial (SDT), ganhou amplitude como política de Estado a partir do Governo do Presidente Luís Inácio Lula da Silva, por intermédio do Ministério do Desenvolvimento Agrário (MDA), especificamente com a criação da SDT, por meio do Decreto nº 5.033, de 5 de abril de 2004<sup>37</sup>. Segundo o *site* da Superintendência de Estudos Econômicos e Sociais da Bahia (SEI),

a Regionalização Territórios de Identidade foi adotada pela Secretaria de Planejamento do Estado da Bahia (Seplan) através da Lei nº 10.705, de 14 de novembro de 2007, quando lançou o Plano Plurianual 2008-2011, e contava à época

<sup>36</sup> Para compreensão maior no que tange à concepção de território e territorialidade, conferir Santos (2014), o qual faz um recorte historicizando tais conceitos, bem como possibilita a compreensão/identificação de seus usos e configurações.

<sup>37</sup> Texto revogado pelo Decreto nº 6.813 de 2009, que o torna sem efeito a partir do Decreto nº 10.473 de 2020.

com 26 Territórios de Identidade que abarcavam os 417 municípios. Desde então, atendendo os objetivos do governo estadual e às necessidades dos municípios baianos sofreu ajustes e atualmente está consolidada conforme Lei nº 13.468, de 29 de dezembro de 2015, que aprova o Plano Plurianual 2016-2019, contando com 27 Regiões. O conceito de Território de Identidade surgiu a partir dos movimentos sociais ligados à agricultura familiar e à reforma agrária, sendo posteriormente adotado pelo Ministério de Desenvolvimento Agrário para a formulação de seu planejamento. A adoção como unidade de gestão para o estado da Bahia foi ajustado conceitual e metodologicamente para a formulação do planejamento em todas as dimensões, incluindo a realidade urbana e as atividades dos setores econômicos secundário e terciário (SEI, 2021).

A definição estabelecida pelo Governo Federal, através do Ministério do Desenvolvimento Agrário, passa a compreender o território não mais apenas como um chão ou uma superfície da terra reconhecida juridicamente. O termo território passa a estabelecer intrinsecamente uma relação com a construção identitária de seu povo, sendo este um lugar representativo de suas trocas sociais, culturais, materiais, espirituais, dentro de

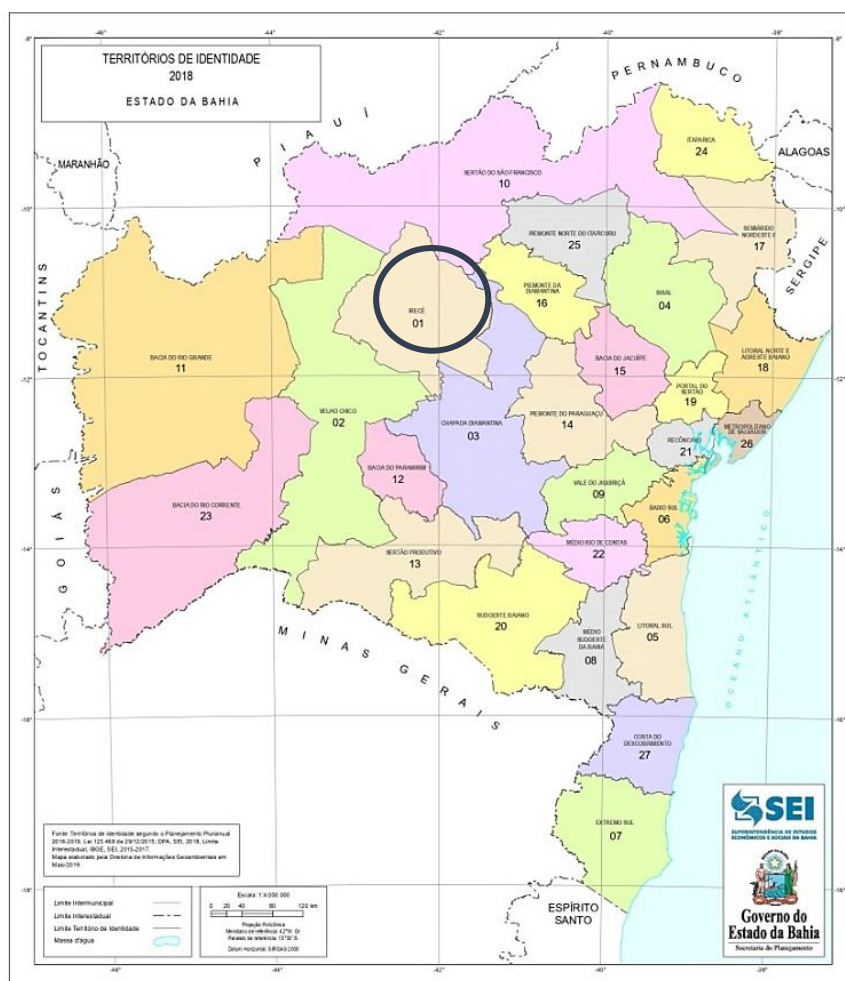
[...] cenários onde os sentimentos de pertencimento e/ou identidade estão atrelados ao que as pessoas sentem em relação aos territórios em que vivem. [...] o território é uma construção coletiva e multidimensional, carregado de múltiplas territorialidades, constituídas sob diversos aspectos e com a participação dos seus atores sociais, superando a ideia de algo pronto, acabado, construído ‘fritamente’, desconsiderando todos os processos anteriores de sua construção. Assim sendo, território deve ser compreendido como construção social, de usos, levando-se em consideração os aspectos históricos, políticos e também de poder presentes neste contexto (Santos, 2014, p. 33).

Como afirma Milton Santos (2007, p. 7), Território é “[...] o lugar em que desembocam todas as ações, todas as paixões, todos os poderes, todas as forças, todas as fraquezas, isto é, onde a história do homem plenamente se realiza a partir das manifestações da sua existência. [...]”. É dessa concepção de Território, de unificação de identidades, que trataremos, tomando-a como *locus* do Projeto ELiHS.

#### 4.1.1 Caracterização sociodemográfica do Território de Identidade de Irecê

Assim, no entrecruzar de realidades locais, o Governo do Estado da Bahia passa a reconhecer a existência de 27 (vinte e sete) Territórios de Identidades, os quais foram organizados e delimitados a partir de suas representações locais, considerando as especificidades de cada região. Vejamos na *Figura 8* os Territórios de Identidades demarcados no Estado da Bahia.

**Figura 8** – Territórios de Identidade demarcados no Estado da Bahia



Fonte: SEI (2018).

A Superintendência de Estudos Econômicos e Sociais da Bahia, através da publicação *Perfil dos Territórios de Identidade da Bahia* (2018), oferece à sociedade informações relativas às características geográficas, socioeconômicas e culturais dos municípios que compõem os Territórios de Identidade na Bahia. Segundo a SEI (2018, p. 148), a maior parte dos municípios que compõem o território apresenta características similares, a saber:

- (i) Baixo nível de urbanização;
- (ii) Nível intermediário de desenvolvimento humano;
- (iii) Elevada participação de pessoas ocupadas na produção para o próprio consumo;
- (iv) Equilíbrio das atividades econômicas na geração de riquezas.

O perfil socioeconômico similar dá ao TII um comportamento de homogeneidade, favorecendo o desenho de um plano de desenvolvimento territorial em que agrega contextos e

fatos que marcam a cultura e a história desses lugares, um novo padrão de regionalização que foi implantado no Estado da Bahia, utilizando o conceito de Territórios de Identidade.

No conceito de território estão agregados os sentimentos de apropriação de uma porção do espaço, assim como, quanto ao seu limite, a sua fronteira [...]. Desta maneira, o limite do território não é necessariamente físico, mas pode se estender até onde a comunidade reconhece a sua influência, o seu exercício de poder [...]. O território é, na sua essência, um fato espacial e social secularmente atrelado a uma dimensão política, permeado de identidade, possível de categorização e de dimensionamento. Nele estão gravadas as referências culturais e simbólicas da população, do grupo ou da comunidade (Anjos, 2011, p. 16).

É nesse cenário de encontro de identidades que compreendemos a concepção de TII, o qual demograficamente compreende uma área de 26.638,40 Km<sup>2</sup>, o que corresponde a aproximadamente 4,7% do território estadual. Geograficamente, localiza-se na região centro-oeste do estado, integrando o Semiárido baiano, fazendo divisa ao oeste com o Território do Velho Chico, ao sul com o Território da Chapada Diamantina e a leste com o Território Piemonte da Diamantina<sup>38</sup>.

**Figura 9** – Composição do Território de Identidade de Irecê – BA



Fonte: Acervo do Projeto ELiHS, mapa elaborado por Santiago (2021).

<sup>38</sup> Dados apresentados pelo Plano Territorial de Desenvolvimento Rural Sustentável e Solidário (PTDRSS, 2017) e Perfil dos Territórios de Identidade da Bahia (2018).

Conforme *Figura 8*, o Território é composto administrativamente por 20 (vinte) municípios que se originaram de desmembramentos dos municípios de Irecê, a saber: América Dourada, Barra do Mendes, Barro Alto, Cafarnaum, Canarana, Central, Gentio do Ouro, Ibipeba, Ibititá, Ipupiara, Irecê, Itaguaçu da Bahia, João Dourado, Jussara, Lapão, Mulungu do Morro, Presidente Dutra, São Gabriel, Uibaí e Xique-Xique. O Plano Territorial de Desenvolvimento Rural Sustentável e Solidário (PTDRSS, 2017, p. 12) ratifica que

Os municípios que integram o Território possuem características sociais, ambientais, econômicas e culturais semelhantes, apresentando áreas de mineração, de exploração da pecuária extensiva, de agricultura de subsistência e, a partir da década de 1970, com o avanço e predomínio da modernização da agricultura com a Revolução Verde, de uma agricultura comercial, capitalizada, com uso intensivo de tecnologias e agroquímicos.

Totalmente inserido no semiárido baiano, integra o Bioma Caatinga, sendo possível distinguir três subsistemas: o Platô de Irecê, com alturas geográficas médias de 700m, ocupando 60% da região e com solos extremamente férteis, mas com limitada oferta de água de chuva e água subterrânea; os vales do Rio Verde e Jacaré, que contornam o platô no lado leste e oeste, com boa oferta de água, tanto superficial como subterrânea, porém com solos fracos e; a Chapada Diamantina, com altitudes até 1000m, com clima mais úmido e frio, mas com solos de baixa fertilidade.

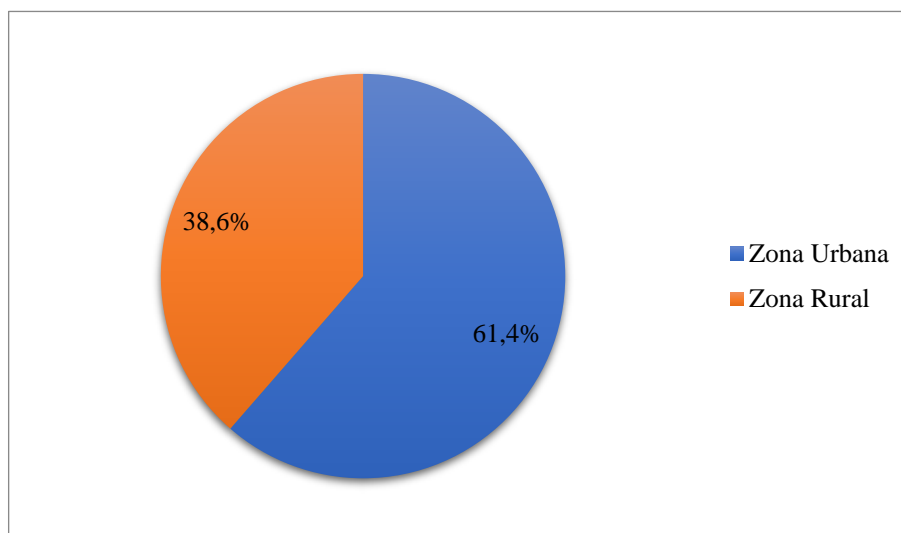
Cabe ressaltar que o processo de povoamento e consolidação do território de Irecê começou em meados do século XIX, tendo como principal pressuposto o início do cultivo da agricultura, mesmo que de forma peculiar, sendo que os 18 municípios que fazem parte desse Território foram oriundos do desmembramento de Morro do Chapéu, Irecê, Central e Xique-Xique. Só assim foram se multiplicando os municípios, cada um com seu modo de produção agrícola, econômica e cultural, até chegar a um produto comum e economicamente rentável, neste caso, o feijão, tornando o território um dos maiores produtores desse grão entre as décadas de 1980 e 1990, situação essa que tornou a cidade conhecida no Estado da Bahia como “terra do feijão” (Ferreira, 2019, p. 20).

O aumento na produção de grãos favoreceu as trocas comerciais entre os municípios, por conseguinte intensificando o processo de ocupação de terras para o cultivo do feijão e outras lavouras na região de Irecê, o que desencadeou o surgimento de novas perspectivas econômicas para os produtos, até então de subsistência, para o comércio local, regional e até estadual. Esse contexto favorece uma caracterização ainda ruralizada ao Território, sendo endossado quando observamos que “em relação ao estrato de moradia, do total de habitantes do território, 61,4% residiam no meio urbano, e 38,6%, no meio rural, perfazendo um grau de urbanização inferior à média do estado, que era de 72,1% em 2010, *caracterizando o TI Irecê como pouco urbanizado.*” (SEI, 2018, p. 148, grifo nosso).



Com uma predominância de habitantes na zona urbana (61,4%), enquanto que, na zona rural, residiam 38,6% do total de habitantes do território, no ano de 2010, podemos afirmar que ainda há um significativo número populacional residente em estratos rurais.

**Gráfico 2** – População censitária (2010) por situação de moradia no TII



Fonte: Elaborado pela autora.

Desse contingente, ainda é possível efetuar uma reanálise mais específica e pontuar, ao menos, 08 (oito) municípios do TII cuja proporção da população na zona rural não ultrapassa 50,0%, reafirmando o caráter pouco urbanizado do Território, especialmente se o compararmos com a média estadual de 72,1% (IBGE, 2010).

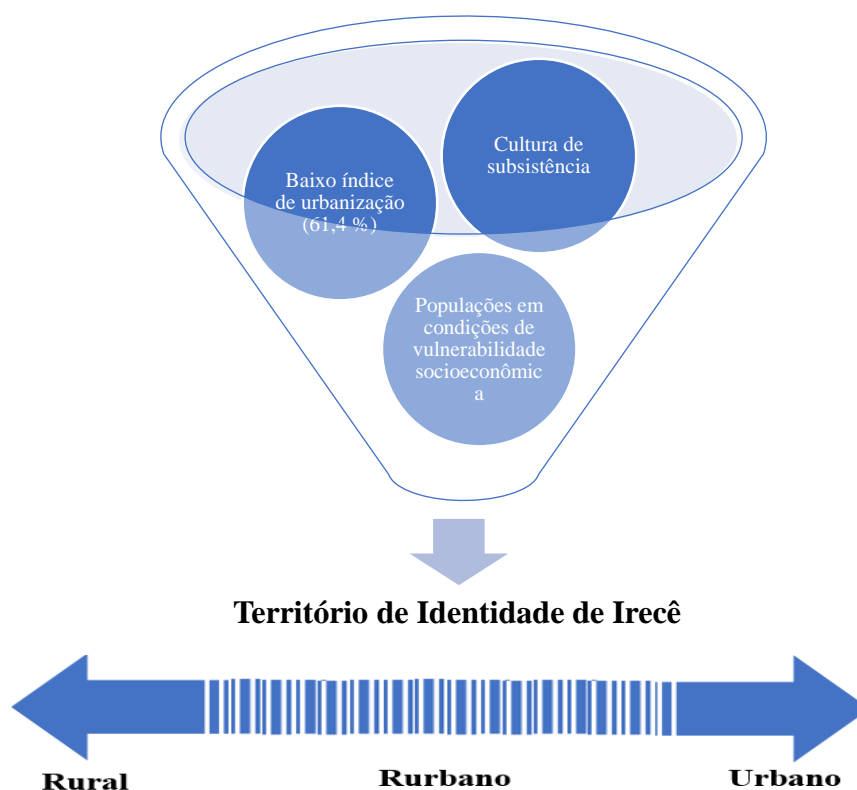
**Gráfico 3** – Grau de urbanização – Bahia, TI Irecê e municípios do TI (2010)



Fonte: SEI (2018).

Cabe ainda frisar, em relação à urbanização no TII, que muitas aglomerações consideradas urbanas são tipicamente rurais em função das atividades agropecuárias nelas desenvolvidas, que são a base de sobrevivência da sua população, visto que o IBGE tem considerado como urbanos pequenos povoados, sedes dos distritos, desde que contenham um mínimo de infraestrutura. Logo, questiona-se: como considerar tais especificidades sociodemográficas? Trata-se de uma definição de urbanização temerária, sem pensar no homem **DO** campo que não vive mais **NO** campo, mas que se desloca diariamente para suas unidades de produção/subsistência. Nesse contexto, cabe refletirmos sobre a possibilidade de fazermos uma releitura da proposta da professora Bortoni-Ricardo (2011) e não pensarmos a ruralização e a urbanização em polos distintos, mas em um *continuum* rural-urbano territorial – nos termos da autora, rurbano.

**Figura 10** – Caracterização do contexto de rurbanização do TII

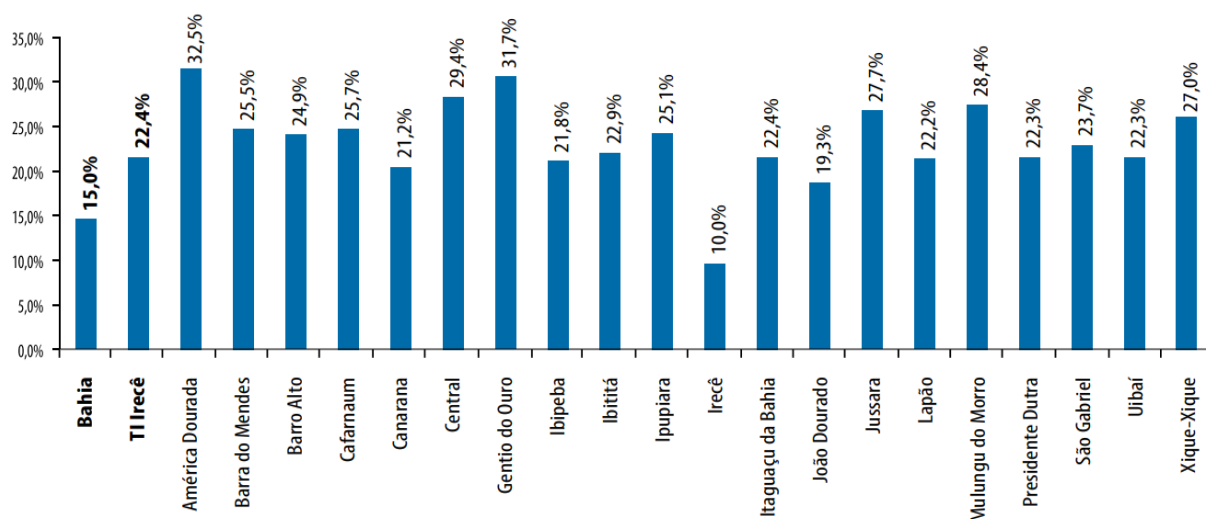


Fonte: Elaborada pela autora.

Um outro importante fator, considerado na *Figura 10*, que endossa a caracterização sociodemográfica do TII, são os grupos de vulnerabilidade social presentes no Território. Para tanto, iniciaremos apresentando o gráfico que marca quantitativamente os índices que retratam

a população que ainda em 2010 se encontrava em extrema pobreza (cf. Gráfico 4) e, em seguida, um retrato dos grupos vulneráveis (cf. Figura 11).

**Gráfico 4** – Proporção de população em extrema pobreza – Bahia, TI Irecê e municípios do TI (2010)

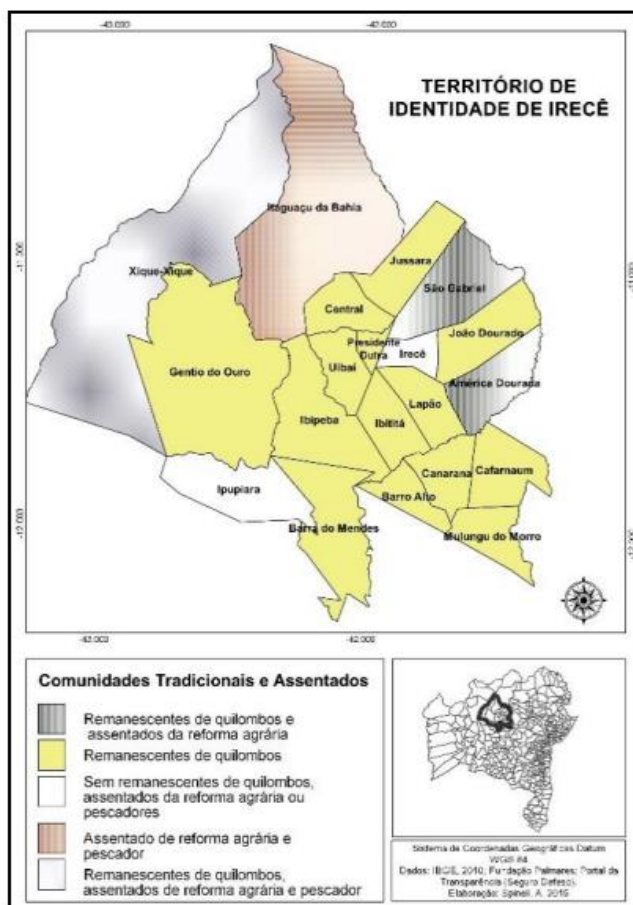


Fonte: SEI (2018).

A leitura do *Gráfico 4* nos permite identificar que, no território, a pobreza extrema era, em 2010, mais elevada em comparação à média estadual: Bahia, 15%, TI Irecê, 22,4%. E o menor indicador, abaixo do da Bahia, é o de Irecê (10%)<sup>39</sup>, consequência da sua caracterização enquanto polo comercial da região. Nesse cenário, surgem os grupos que se encontram em condições de vulnerabilidade social, quais sejam: os assentados da reforma agrária, pescadores e remanescentes de quilombos.

<sup>39</sup> O índice verificado no município de Irecê teve um impacto positivo no índice do território, embora a maioria dos municípios tenha apresentado uma elevada proporção de sua população vivendo em extrema pobreza.

**Figura 11** – Grupos vulneráveis no TII



Fonte: PTDRS (2017 *apud* IBGE, 2010; INCRA, 2015; Fundação Palmares, 2015)

Dentre as famílias que se encontram em situação de vulnerabilidade socioeconômica decorrente da pobreza, do precário ou nulo acesso aos serviços públicos, estão as famílias de remanescentes quilombolas, tratadas nesta tese enquanto comunidades afro-brasileiras, as quais ganham nas seções que seguem uma descrição cuidadosa, do mapeamento à caracterização das comunidades hoje presentes no TII.

#### 4.1.2 Dos Quilombos às Comunidades rurais afro-brasileiras: da descrição terminológica ao enquadramento histórico

Pelo fato de o termo quilombo possuir na literatura especializada uma significação tão ampla, apresentaremos nesta seção alguns conceitos já estabelecidos, bem como a proposta assumida nesta tese de compreendermos essas comunidades como *afro-brasileiras*, por acreditarmos que linguisticamente estaremos mais próximos do que conceitualmente entende-se por remanescentes de quilombos no século XXI, considerando a sua constituição atual e as suas trajetórias ancestrais. Sendo assim, não temos a pretensão de descrever o fenômeno das

origens do termo *quilombo*, mas, sim, buscar uma terminologia que esteja direcionada aos aspectos sociais, culturais, históricos e que se encaixe na proposta linguística do Projeto ELiHS.

A historiografia do quilombo marca as lutas e as conquistas pela permanência e reconhecimento de seus territórios ancestrais, e, nesse contexto, Nascimento (2018b, p. 280-281, grifo da autora) afirma que

A primeira referência a quilombo que surge em documento oficial português data de 1559, mas somente em 2 de dezembro de 1740 as autoridades portuguesas definem ao seu modo o significado do quilombo. O nordeste brasileiro emergia de um século de guerras em que vários quilombos, sobretudo aqueles reunidos na República de Palmares, resistiam às investidas militares contra os núcleos de população negra livre do domínio colonial. Depois dessas guerras, da destruição de Palmares e da agitação que se processou ao redor desse núcleo no século XVII, houve um recrudescimento da resistência de outros quilombos. Foi nesse momento que as autoridades portuguesas, assustadas, definiram o quilombo como sendo: *‘Toda a habitação de negros fugidos que passem de cinco, em parte desprovida, ainda que não tenham ranchos levantados nem se achem pilões neles’*.

Ainda segundo Nascimento (2018b, p. 289), “no final do século XIX, o quilombo recebe o significado de instrumento ideológico contra as formas de opressão [...]. Esta passagem de instituição em si para símbolo de resistência mais uma vez redefine o quilombo”. Assim, a autora segue afirmando que é como caracterização ideológica que o quilombo inaugura o século XX, visto que deixa de ser visto apenas como resistência à escravidão e passa “[...] a ser sinônimo de povo negro, de comportamento do africano e de seus descendentes e esperança para uma melhor sociedade” (Nascimento, 2018b, p. 292). Sendo assim, o quilombo passa a ser

uma condição social, fundamentalmente uma condição social, quer dizer, ele não se esgota no militarismo, na guerra que foi feita em relação aquela que ele reagiu, mas a estrutura do quilombo, o que realmente singulariza o quilombo, é que ele é um agrupamento de negros, que o negro empreende, que aceita o índio dentro dessa estrutura e *que não foi aceito nunca dentro da sociedade brasileira, como ainda não é aceito até agora [...]* (Nascimento, 2018a, p. 126, grifo nosso).

Aqui, compreendemos que a estrutura quilombo sofreu modificações ao longo da História, mas a principal característica que se mantém é a de ser uma organização social empreendida pelos negros (Nascimento, 2018a, p. 129), o que nos faz reafirmar o conceito definido pelo professor Rafael Sanzio Araújo dos Anjos (2011, p. 17, grifo do autor) de que

o quilombo era uma reconstrução e elaboração concreta de um tipo de organização territorial africana no ‘novo espaço’ denominado Brasil. Durante os quase quatro séculos de tensões e confrontos de classes no sistema escravista, os quilombos funcionaram como uma verdadeira ‘válvula de escape’ para diluir a violência da escravidão. A palavra *quilombo*, que tem sua origem na língua banto e possui referência em expressões como: habitação, acompanhamento, albino, floresta e guerreiro.

É nesse contexto que, a partir da Constituição Brasileira de 1988, o termo quilombo adquire uma nova significação, ao ser inscrito no art. 68 do Ato das Disposições Constitucionais Transitórias (ADCT): “Aos remanescentes das comunidades de quilombos que estejam ocupando suas terras, é reconhecida a propriedade definitiva, devendo o Estado emitir-lhes títulos respectivos” (BRASIL, 1988). Frente a isso, as comunidades se mobilizaram, a fim de serem reconhecidas Brasil afora enquanto *Remanescentes quilombolas*.

No entanto surge o primeiro questionamento: quem são os chamados remanescentes de quilombos cujos direitos são atribuídos pela Constituição? Do ponto de vista jurídico e para fins de políticas sociais destinadas aos quilombolas brasileiros, temos uma definição normativa através do Decreto nº 4.887/2003, em seu artigo 2º: “Consideram-se remanescentes das comunidades dos quilombos, para os fins deste Decreto, os grupos étnico-raciais, segundo critérios de auto atribuição, com trajetória histórica própria, dotados de relações territoriais específicas, com presunção de ancestralidade negra relacionada com a resistência à opressão histórica sofrida” (BRASIL, 2003).

A definição jurídica do termo *remanescente de quilombo* imprime discussões em torno do aparato conceitual da categoria *quilombo*, levando pesquisadores das Ciências Humanas e Sociais a buscarem alternativas para uma questão que até hoje desperta dúvidas e incertezas, fazendo-se necessário este momento de reflexão para acertadamente enquadrarmos o termo assumido, nesta tese, de *comunidades afro-brasileiras*.

Vejamos que Gusmão (1993) afirma que o que é posto pela Constituição esbarra em questões que precisam urgente serem desfeitas, visto que existem vazios que vão desde o desconhecimento do que seja ser negro brasileiro até a conceituação do que seja “quilombo”. Assim, a autora afirma que

[...] a questão pois, de continuidade histórica posta pela condição remanescente, esbarra na sedução política e ideológica de como foi pensado o ‘alargamento’ do conceito de quilombo e exige a superação de equívocos quanto ‘as expectativas usuais de senso comum face ao processo de definição de uma identidade étnica. Existir uma identidade que unifique e singularize uma população não é de maneira alguma uma necessidade cultural...’, diz Oliveira (op.cit. :123) mas exige a compreensão histórica de definição de território e de limites de grupo frente ao confronto com outros grupos (Gusmão, 1993, p. 11).

Nos termos de Gusmão (1993), ser remanescente significa “o que não morreu”, “o que sobreviveu”, os que são “vivos” e não sobreviventes. Sendo assim, “[...] a questão da continuidade histórica, pensada na condição de remanescente de quilombo, pode vir-á-ser uma perigosa contra-prova aos interesses dos grupos [...]” (Gusmão, 1993, p. 10). Enfim, “[...] esse quilombo hoje não pode ser entendido historicamente como esse quilombo de ontem. O

*quilombo de hoje significa muito mais uma consciência, uma ideologia realmente, uma consciência de que você é um homem, que você é capaz de empreender coisas capazes de sempre aceitas, viver e ser aceito de uma sociedade, é isso que é fundamental”* (Nascimento, 2018a, p. 137, grifo nosso).

Nesse sentido, a Associação Brasileira de Antropologia (ABA), em outubro de 1994, reuniu-se ao Grupo de Trabalho sobre Comunidades Negras Rurais na tentativa de elaborar um conceito melhor delineado para o termo “remanescente de quilombo”:

o documento procurou desfazer os equívocos referentes à suposta condição remanescente, ao afirmar que ‘contemporaneamente, portanto, o termo não se referia a resíduos arqueológicos de ocupação temporal ou de comprovação biológica’. Tratava-se de desfazer a ideia de isolamento e de população homogênea ou como decorrente de processos insurrecionais. O documento posicionava-se criticamente em relação a uma visão estática do quilombo, evidenciando seu aspecto contemporâneo, organizacional, relacional e dinâmico, bem como a variabilidade das experiências capazes de serem amplamente abarcadas pela ressemantização do quilombo na atualidade. Ou seja, mais do que uma realidade inequívoca, o quilombo deveria ser pensado como um conceito que abarca uma experiência historicamente situada na formação social brasileira.

A conceituação de quilombo do documento da ABA ampliou a visão do fenômeno referido e conferiu-lhe uma maior pertinência em relação aos pleitos já formulados. Faltava identificar o sujeito do direito, os critérios normativos para a regulamentação da lei e sua aplicabilidade, ou seja, os procedimentos e etapas a serem cumpridos para a titulação das terras, as responsabilidades e competências dos atores sociais envolvidos. Ao contrário do que se supunha, a questão, a partir daí, revelou-se de grande complexidade, pois tratava-se de considerar não apenas os aspectos referentes à identidade dos negros no Brasil, mas os vários atores envolvidos e os inúmeros interesses conflitantes sobre o patrimônio material e cultural brasileiro, ou seja, questões de fundo envolvendo identidade cultural e política das minorias de poder no Brasil (Leite, 2000, p. 341-342).

É nesse contexto conceitual, do campo histórico ao antropológico, que recai a questão incitada nesta seção de definirmos as 148 comunidades presente no TII, tentando, sempre que possível, “[...] compreender a problemática sob três vertentes, a do texto, do contexto e do pretexto; o trocadilho é intencional, afinal, é assim toda escritura que tem por alvo um público determinado” (Boletim informativo, 1997, p. 6), em nosso caso, a comunidade externa, a quem se direcionam os interesses históricos e culturais, e a comunidade linguística, a quem se direcionam os interesses sociolinguísticos. Dois grandes campos de atuação e o ELiHS, um projeto transdisciplinar.

Segundo Leite (2000), a partir da transição do século XIX para o XX, é comum nos depararmos em textos científicos com a expressão “quilombolas”, “remanescentes de quilombos”, “pretos” ou “afro-brasileiros”, indicando um recorte de grupo que vem lutando pela valorização identitária e que é perversamente invisibilizado pelas e nas esferas públicas. Nesse contexto, ainda terminológico, Anjos (2011, p. 187, grifo do autor) afirma que

No Brasil, as comunidades negras tradicionais, os remanescentes de quilombos, mocambos, comunidades negras rurais, quilombos contemporâneos, comunidades quilombolas ou terras de preto, referem-se a um mesmo patrimônio territorial e cultural inestimável [...]. O conceito de comunidade quilombola tem uma referência no campesinato *negro*, de povo de matriz africana que conseguiram ocupar uma terra e manter uma autonomia política e econômica. Ao quilombo contemporâneo está associada uma interpretação mais ampla, não somente de resistência no passado, mas sobretudo, no presente. Um território étnico capaz de se organizar e se reproduzir no espaço geográfico de condições adversas, ao longo do tempo e com resistência para a manutenção da sua forma particular de viver.

Como afirma Anjos (2011), hoje, no Brasil, os remanescentes de antigos quilombos, comunidades negras rurais, quilombos contemporâneos, comunidades quilombolas referem-se a um mesmo patrimônio territorial e cultural e, em grande parte, desconhecido pelo Estado. A história brasileira tem, muitas vezes, se referido a essas comunidades sempre no passado, como se não fizessem mais parte da vida do país, no entanto torna-se urgente compreendermos que

os territórios quilombolas atuais estão inseridos no bojo das populações tradicionais brasileiras, que constituem grupo de grande relevância para configuração da identidade nacionais [...] os seus sítios localizados atualmente nos espaços, rural e urbano no Brasil, constituem territórios étnicos de resistência secular, de identidade marcante, de resgate histórico e de manutenção das heranças africanas sobreviventes. Não podemos perder de vista que, nem todos os territórios quilombolas existentes no Brasil, que se enquadram nesse conceito mais *largo*, são *remanescentes* dos antigos quilombos e os que porventura forem, muitas vezes não terão como provar sua historicidade (Anjos, 2011, p. 18-19, grifo do autor).

De acordo com Gusmão (1992), as terras que hoje constituem as Comunidades de remanescentes quilombolas, comunidades negras rurais ou, como assumiremos aqui, *comunidades afro-brasileiras* que são ocupadas por grupos negros muitas vezes foram doadas por antigos senhores a escravos fiéis, ou então doadas aos santos de devoção, outras vezes foram simplesmente ocupadas por famílias negras após a Abolição, não se enquadrando nesse sentido historiograficamente convencional/tradicional.

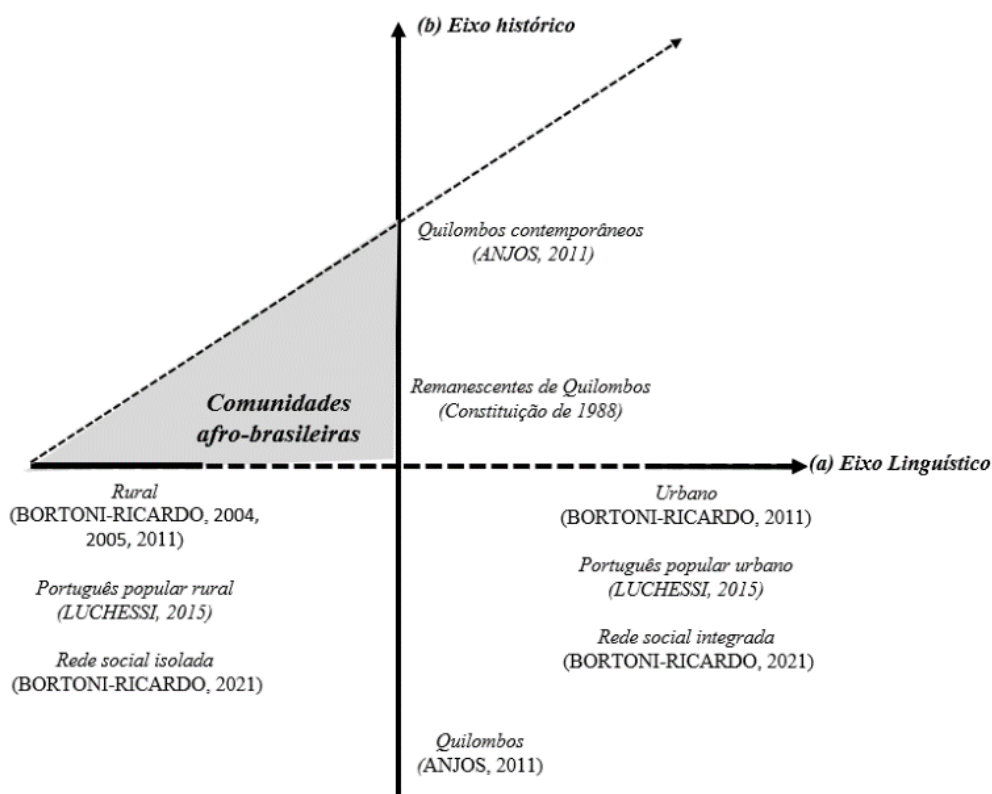
Nesse contexto, entendemos que mudanças terminológicas implicam ordens preestabelecidas na forma de discurso, na reprodução de novos conceitos, novas ideologias; por isso, ao assumirmos nesta tese o termo *comunidades afro-brasileiras*, consideramos que

o conceito de afrodescendência implica que a matriz africana foi recebida e desenvolvida no Brasil. A africanidade foi transportada através do oceano atlântico e desenvolvida no novo território dando respostas as condições geográficas, sociais, políticas e econômicas novas *Afrodescendência implica na existência de um patrimônio cultural, intelectual e civilizatório desenvolvido no Brasil, tendo como base o capital cultural diverso africano, sem, contudo, perder os eixos de força da africanidade* (Cunha Júnior, 2022, p. 106, grifo nosso).



Acreditamos que, a partir dos trabalhos citados até o momento, assumir o termo *comunidades afro-brasileiras* afasta-nos de uma visão eurocêntrica e conseguimos trazer para o centro da discussão a importância desses povos na formação histórica do Brasil, assim comungando com o campo linguístico em que Lucchesi (2015, 2009b; 2003), em razão de um aprendizado irregular (cf. seção 2.3.3), caracteriza o termo *português afro-brasileiro* como uma variedade popular do português falado no Brasil, com uma maior variação no uso de alguns mecanismos gramaticais se comparado às outras variedades desse português, termo que condiz com as amostras de fala do Projeto ELiHS.

**Figura 12** – Do conceito de Quilombo à comunidade afro-brasileira



Fonte: Elaborada pela autora.

A *Figura 12* é uma tentativa gráfica, como já citado, de “[...] compreender a problemática sob três vertentes, a do texto, do contexto e do pretexto; o trocadilho é intencional, afinal, é assim toda escritura que tem por alvo um público determinado” (Boletim informativo, 1997, p. 6). O eixo (b), que verticaliza o eixo (a), representa as terminologias utilizadas pela historiografia, permitindo-nos englobar no termo *comunidade afro-brasileira* as discussões amplificadas que vão desde os *remanescentes de quilombo* até o conceito de *quilombos*

*contemporâneos* (Anjos, 2011). Enquanto o eixo (a), horizontal, é marcado por uma linha tracejada que representa os *continuum* linguísticos, a saber:

- (i) *Rural – Rurbano – Urbano* (Bortoni-Ricardo, 2004, 2005, 2011);
- (ii) *Português afro-brasileiro – Português rural – Português popular rurbano do interior – Português urbano do interior – Português popular rurbano das grandes cidades – Português urbano das grandes cidades* (Luchesi, 2015, p. 218);
- (iii) *Redes isoladas – Redes integradas* (Bortoni-Ricardo, 2021).

A gradação proposta pelos *continuum*, resultado de estudos científicos dos referidos autores, demonstra que, diante das mudanças ocorridas na sociedade contemporânea, consequente diminuição das fronteiras geográficas e sociais, é preciso deixarmos de pensar de forma tão bipolarizada, e que hoje o que Lucchesi, Baxter e Ribeiro (2009) determinam como “comunidades rurais afro-brasileiras isoladas” sofreu também mudanças em suas configurações, não se caracterizando mais como totalmente isoladas. Sendo assim, assumimos aqui o *continuum* (iii), em que compreendemos essas Comunidades dentro de redes mais isoladas (restrita à família extensa e vizinhos) e redes mais integradas (relação com pessoas em muitos contextos sociais), visto que as redes isoladas favorecem a manutenção da cultura rural e consequente focalização do vernáculo, à medida que o migrante de origem rural estabelece um maior número de vínculos em seu novo ambiente, passando a desempenhar novos papéis sociais, e suas redes passam a ser consideradas mais integradas, com consequente difusão do vernáculo (Bortoni-Ricardo, 2021, p. 74).

Assim, o Projeto ELiHS, que entrelaça aspetos linguísticos e históricos, assume a terminologia *comunidades afro-brasileiras*, assim como Luchesi, Baxter, Ribeiro (2009) e Santana, Araújo e Freitag (2018a, 2018b), dentre outros que têm assumido essa terminologia no campo da linguística. Desse modo, ainda de forma muito localizada, pretendemos tornar as 148 comunidade afro-brasileiras, situadas no TII, visíveis perante ao mapa sociolinguístico do português brasileiro, trazendo-as não apenas como um registro histórico, reduzidos em um espaço/tempo pretérito, mas projetando-as no presente, em um estudo linguisticamente sincrônico, do presente para o passado.

#### 4.1.3 Comunidades rurais afro-brasileiras no TII: do mapeamento cartográfico à visibilidade territorial

Rafael Anjos, em 1999, já apontava que, dentre os pontos estruturais que permeiam a situação dos descendentes dos antigos quilombos no Brasil, destaca-se a carência de informações sistematizadas referentes à distribuição dessas comunidades no território. Infelizmente, “as estimativas são inconsistentes, divergentes e não existem pesquisas direcionadas para investigar a questão com essa abordagem geográfica” (Anjos, 1999, p. 9-10), o que nos traz um alerta para urgência de reunirmos os dados que colaborem para caracterização, ainda que generalizada, das comunidades que carregam em sua formação fragmentos que compõem a história da constituição do povo brasileiro e da língua oficializada, impositivamente, enquanto Português.

O Projeto ELiHS, sediado na UNEB – *Campus* de Irecê, enxerga o inexistir de informações sobre as 148 comunidades certificadas, pela Fundação Cultural Palmares, enquanto Comunidade de Remanescentes de Quilombo como um desafio e um convite à pesquisa científica, na tentativa de organizar material empírico que colabore para caracterização da identidade linguística desses povos, compreendendo-a como um patrimônio imaterial que precisa ser preservado, guardado, por ser algo que corre o risco de ser “apagado” diante do crescente processo de urbanização, proporcionado pelo contexto do século XXI.

Hoje ainda é possível notar que, mesmo transcorridos mais de um século desde a abolição do escravismo, a maior parte da população negra brasileira continua concentrada nas regiões menos desenvolvidas do país, e, quando pensamos nas comunidades afro-brasileiras, muitas dessas ainda se encontram em estado de vulnerabilidade social e “invisíveis” à sociedade e aos poderes públicos. Sim, essa é a realidade das comunidades situadas no Território de Irecê, mesmo que os membros representativos pouco ou nada tenham registrado, ainda que extraoficialmente, sobre suas comunidades, o que torna salutar a tentativa de resgatamos os fragmentos de informações que possam permitir uma compreensão mais aproximativa da realidade sociodemográfica dessas comunidades.

Portanto, na tentativa de colaborar no preenchimento dessas lacunas seculares, aqui nos propomos a apresentar dados informacionais inéditos que caracterizarão, de forma ainda generalizada, as comunidades afro-brasileiras situadas no TII. As informações a serem apresentadas apoiam-se nos levantamentos realizados entre os anos de 2020-2021, conforme os seguintes procedimentos metodológicos: Levantamento das comunidades certificadas pela FCP; Rede de contatos com membros representativos das Comunidades; Envio de formulário eletrônico.

Diante do levantamento das comunidades existentes no TII, chegamos a um quantitativo de 148 comunidades certificadas, desenhando uma missão desafiadora de estabelecer contato telefônico com membros que fossem representativos e aptos a responderem as questões propostas no formulário eletrônico (via *Google Forms*<sup>40</sup>). Vejamos as perguntas que conduziram os resultados quantitativos e representativos a serem apresentados nesta seção:

**Quadro 7** – Questões presentes no formulário eletrônico

	DESCRIÇÃO DAS QUESTÕES	FORMA DE RESPOSTA
Q. 1	Nome completo	Aberta
Q. 2	Qual a função que você exerce na Comunidade?	Aberta
Q. 3	Qual o nome da Comunidade Quilombola?	Aberta
Q. 4	A qual cidade pertence a Comunidade?	Múltipla escolha
Q. 5	Já deram entrada ao processo de reconhecimento pela Fundação Palmares?	Múltipla escolha (Sim / Não)
Q. 6	Já deram entrada ao processo de regularização do Incra?	Múltipla escolha (Sim / Não)
Q. 7	Já participaram de algum projeto apoiado pela Fundação Palmares?	Múltipla escolha (Sim / Não)
Q. 8	A Comunidade possui Associação?	Múltipla escolha (Sim / Não)
Q. 9	Sua Comunidade tem, em média, quantos anos de existência/formação?	Aberta
Q. 10	A Comunidade possui, em média, quantos habitantes/membros?	Aberta
Q. 11	As pessoas, da Comunidade, consideradas negras representam?	Múltipla escolha (100%, 90%, 80%, 70%, 60%, 50%, Menos de 50%)
Q. 12	Os membros da comunidade são em sua maioria:	Múltipla escolha (Idosos (acima de 60 anos); Adultos (entre 31 e 59 anos); Jovens (entre 18 e 30 anos); Adolescentes/Crianças (menores de 18 anos)).
Q. 13	Qual a religião predominante na Comunidade?	Múltipla escolha (Matriz africana (Candomblé e Umbanda); Católica; Evangélica; Outra).
Q. 14	A Comunidade possui acesso telefônico?	Múltipla escolha (Sim, com fácil acesso; Sim, com difícil acesso; Não).

<sup>40</sup> O formulário pode ser acessado através do link [https://docs.google.com/forms/d/e/1FAIpQLScO\\_H5DuA5axcfcpKKaVkJdHQwwfOhREsjwd5rZIYjubsZcwg/viewform?usp=sf\\_link](https://docs.google.com/forms/d/e/1FAIpQLScO_H5DuA5axcfcpKKaVkJdHQwwfOhREsjwd5rZIYjubsZcwg/viewform?usp=sf_link).

Q. 15	Em relação à internet:	Múltipla escolha (NÃO possuem acesso à internet; 100% dos moradores possuem acesso à internet; Entre 60% e 90% dos moradores possuem acesso à internet; 50% dos moradores possuem acesso à internet; Menos 50% dos moradores possuem acesso à internet).
Q. 16	A Escola da Comunidade atende a quais níveis de ensino? (Caso necessário, marque mais de uma alternativa).	Múltipla escolha (Não possui escola na Comunidade; Apenas Creche; Educação Infantil (4 a 5 anos de idade); Educação Fundamental I (1º ao 5º ano); Turma Multisseriada (várias séries); Educação Fundamental II (6º ao 9º ano); Ensino Médio).
Q. 17	Qual o percentual de pessoas/membros que concluíram o Ensino Médio?	Múltipla escolha (Menos 50% concluíram o Ensino Médio; 50% concluíram o Ensino Médio; Entre 60% e 90% concluíram o Ensino Médio; 100% concluíram o Ensino Médio).
Q. 18	A comunidade fica a quantos quilômetros da sede?	Aberta
Q. 19	Quanto ao acesso à sede do município, você considera as condições da estrada:	Múltipla escolha (Boa; Regular; Ruim)
Q. 20	A Comunidade possui algum tipo de documento/material que narre sua história/formação?	Aberta

Fonte: Elaborada pela autora.

Tal atividade não seria possível sem a criação de uma rede de apoio que iniciou com auxílio de estudantes quilombolas da UNEB/*Campus XVI*, dos bolsistas de extensão e da parceria com o CODETER (Conselho de Desenvolvimento Territorial), que possibilitou uma rede de articulação entre a pesquisadora e alguns representantes quilombolas do Território. Como afirma Anjos (1999, p. 11, grifo nosso),

é até sintomático que a distribuição dos remanescentes de quilombos no Brasil não haja merecido estudos de conjunto referentes a sua espacialização, seja por questões políticas, pela abrangência interdisciplinar, *pelas dificuldades de acesso e levantamento dos dados*, seja pela necessidade de utilização de tecnologias e ferramentas sofisticadas com grande capacidade de integração de dados ou mesmo pelo desafio de tratar a temática com uma lente que permita uma visão do todo.

Diante do cenário de invisibilidade que afeta as comunidades situadas no TII, as quais não possuem registro algum nas Secretarias Municipais, fomos inquietados a comungar aos objetivos linguísticos desta tese a caracterização geral dessas comunidades, que está disponível no *site* do Projeto ELiHS (*cf.* seção 4.3), tornando acessíveis ao público dados inéditos que, ainda de forma incipiente, têm mobilizado a população civil e grupos políticos a perceberem a existência e resistência das Comunidades **no** e **do** TII.

De início, faz-se importante a apresentação do mapeamento realizado, na tentativa de uma representação cartográfica dessas Comunidades, ressaltando que a cartografia “[...] é uma ferramenta com condições concretas de representar o que está acontecendo no território, bem como de apontar indicadores geográficos para o processo de reconhecimento e de delimitação de um sítio quilombola” (Anjos, 1999, p. 88).

**Figura 13** – Mapa das Comunidades certificadas pela FCP enquanto CRQ



Fonte: Dados da autora, mapa elaborado por Santiago (2021).

Para disponibilização e facilitação no processo de leitura das comunidades já mapeadas no território, organizamos um quadro descritivo em que constam os atuais<sup>41</sup> nomes das comunidades, suas respectivas cidades sedes, processos de Declaração enquanto Remanescentes Quilombolas, pela Fundação Cultural Palmares, e, por fim, processo de abertura para regularização de terras pelo Instituto Nacional de Colonização e Reforma Agrária (doravante INCRA).

**Quadro 8** – Mapeamento das Comunidades certificadas pela FCP enquanto CRQ

CIDADES	CRQ	Nº do Proc. FCP	Nº do Proc. INCRA
<b>AMÉRICA DOURADA</b>	Alegre	01420.003594/2013-55	-
	Barriguda dos Bidós	01420.003592/2013-66	-
	Boa Esperança	01420.003082/2013-99	-
	Boa Vista	01420.003593/2013-19	-
	Campo Alegre	01420.003591/2013-11	-
	Canabrava	01420.007499/2010-88	-
	Garapa	01420.007495/2010-08	-
	Lagoa Verde	01420.006570/2010-13	-
	Lagoa Verde de Baía	01420.006570/2010-13	-
	Queimada Nova dos Vianas	01420.006570/2010-13	-
	Lajedão dos Mateus	01420.003381/2009-47	-
	Lapinha	01420.000173/2010-20	54160.002239/2013-33
	Mulungu	01420.003590/2013-77	-
	Queimada dos Beneditos	01420.007495/2010-08	54160.002075/2013-44
	Sarandí	01420.003587/2013-53	-
Prevenido	01420.003584/2013-10	-	
<b>BARRA DO MENDES</b>	Antarí	01420.003580/2013-31	-
	Canarina	01420.003579/2013-15	-
	Queimada do Rufino	01420.003582/2013-21	-

<sup>41</sup> Muitas comunidades recentemente assumiram novos nomes, para autoidentificação, o que envolve outras tantas questões vinculadas a preconceito e exclusão, a saber: de *Lagoa dos Negros* para *Lagoa do Gaudêncio*. Um convite a estudos toponímicos, numa interface entre língua, cultura, história e geografia.

	Volta do Ábdon	01420.003581/2013-86	-
<b>BARRO ALTO</b>	Barreirinho	01420.000630/2008-61	-
	Malvinas	01420.002832/2008-48	-
	Segredo	01420.002141/2008-44	-
	Rua do Juá*	01420.002830/2008-59	--
	Volta Grande	01420.002828/2008-80	-
<b>CANARANA</b>	Brejinho	01420.001541/2006-71	54160.000408/2012-10
	Cansanção	01420.010020/2011-71	54160.001939/2014-91
	Cruzeiro	01420.010016/2011-11	-
	Floresta I	01420.010018/2011-01	-
	Floresta II	01420.010017/2011-58	54160.001935/2014-11
	Lagoa do Zeca	01420.001542/2006-15	54160.002845/2011-97
	Largo do Miranda	01420.010015/2011-69	54160.002692/201340
	Mato Verde	01420.010022/2011-61	-
	Novo Horizonte	01420.010023/2011-13	54160.001936/2014-58
	Segredo	01420.010019/2011-47	-
	Volta do Angico	01420.010021/2011-16	54160.001937/2014-01
<b>CAFARNAUM</b>	Erva Cidreira	01420.004895/2015-68	-
	Queimado do Tiano	01420.004895/2015-68	-
	Lagoa do Gado	01420.004895/2015-68	-
	Presídio	01420.004894/2015-13	-
	Charel	01420.004894/2015-13	-
	Curralinho	01420.004894/2015-13	-
	Roça do Meio	01420.004894/2015-13	-
	Umbuzeiro	01420.004893/2015-79	-
	Bandeira	01420.004893/2015-79	-
	Salaminho	01420.004893/2015-79	-
<b>CENTRAL</b>	Caldeirãozinho	01420.016242/2013-60	54160.001984/2015-27
	Capoeira da Serra	01420.016756/2013-15	54160.001983/2015-82
	Caroá	01420.016754/2013-26	54160.001990/2015-84
	Floresta	01420.016752/2013-37	54160.001987/2015-61
	Lagoa do Martinho	01420.016749/2013-13	54160.001988/2015-13



	Mandacarus	01420.016747/2013-24	54160.0019892015-50
	Milho Verde	01420.016746/2013-80	54160.001991/2015-29
	Morro de Lúcio	01420.016742/2013-00	54160.001985/2015-71
	São João de Zé de Preta	01420.016740/2013-11	54160.002308/2015-71
	Vereda	01420.016738/2013-33	54160.001986/2015-16
<b>GENTIO DO OURO</b>	Barreiro Preto	01420.000136/2006-35	54000.075106/2020-48
	Mato Grosso	01420.001705/2015-51	54000.075173/2020-62
	Alagoinhas	01420.000959/2010-47	54000.075178/2020-95
	Pacheco	01420.102375/2018-62	54000.075187/2020-86
	Silvério	01420.102373/2018-73	54000.075150/2020-58
	Gregório	01420.102373/2018-73	54000.075150/2020-58
	Olho d'Água	01420.102374/2018-18	54000.075196/2020-77
	Malhada	01420.102374/2018-18	54000.075196/2020-77
	Água doce	01420.102372/2018-29	54000.075192/2020-99
<b>IBIPEBA</b>	Olhos D'Água do Badú	01420.001166/2007-40	-
	Aleixo	01420.016377/2014-14	-
	Lagoa do Cedro	01420.101301/2019-90	-
	Olhos d'Água dos Batatas	01420.016373/2014-28	-
	Lagedinho	01420.100218/2020-37	-
	Mocobeu	01420.103024/2018-79	-
	Salva Vidas	01420.001168/2007-39	-
	São Vicente	01420.016375/2014-17	-
	Segredo	01420.001167/2007-94	-
	Serra Grande	01420.015360/2011-99	-
	Muribeca	01420.016374/2014-72	-
	Abade	01420.016374/2014-72	-
	Santo Antônio	01420.016374/2014-72	-
<b>IBITITÁ</b>	Bairro da Mata	01420.011298/2011-66	54160.001288/2012-78
	Barro Duro	01420.002650/2009-58	54160.000844/2012-99
	Batata	01420.002649/2009-23	54160.000842/2012-08
	Canoão	01420.001339/2009-91	54160.000832/2012-64
	Faveleiro	01420.011293/2011-33	54160.000841/2012-55

	Lagedão	01420.011295/2011-22	54160.000843/2012-44
	Lagoa da Pedra	01420.002651/2009-01	54160.001277/2012-98
	Pedra Lisa	01420.001224/2010-31	54160.000833/2012-17
<b>IRECÊ</b>	Baixão de Zé Preto	01420.004357/2017-35	-
	Lagoa Nova	01420.006140/2015-06	-
<b>ITAGUAÇU DA BAHIA</b>	Alegre	01420.002908/2005-92	54160.003353/2008-13
	Barreiros	01420.002908/2005-92	54160.003353/2008-13
	Alto Bebedouro	01420.002380/2008-02	54160.003353/2008-13
	Cajueiro	01420.003065/2009-75	54160.003353/2008-13
<b>JOÃO DOURADO</b>	Angicão	01420.011296/2011-77	54160.001283/2012-45
	Baixa das Cabaças	01420.011302/2011-96	54160.001287/2012-23
	Descoberta	01420.003821/2011-81	54160.001279/2012-87
	Feitosa	01420.001232/2006-09	54160.001281/2012-56
	Lagoa do Barro	01420.013515/2011-52	-
	Lagoa do Meio	01420.003822/2011-25	54160.001280/2012-10
	Lagoa do Rozeno	01420.011284/2011-42	-
	Lagoa dos Lundus	01420.003820/2011-36	54160.001286/2012-89
	Mata do Milho	01420.006550/2011-15	-
	Para Quedas	01420.003818/2011-67	54160.000848/2012-77
	Riacho	01420.006551/2011-60	54160.003751/2011-35
	Sabino	01420.001231/2006-56	54160.001282/2012-09
	Salinas	01420.011294/2011-88	-
	Serra Azul	01420.011287/2011-86	54160.001285/2012-34
	Serrinha	01420.003819/2011-10	54160.001278/2012-32
Sertão Bonito	01420.011288/2011-21	54160.001284/2012-90	
<b>JUSSARA</b>	Algodões	01420.000546/2007-67	-
	Sítio Novo	01420.000118/2007-34	-
<b>LAPÃO</b>	Babilônia	01420.010061/2011-68	-
	Lagoa do Gaudêncio	01420.002159/2009-27	54160.001065/2010-49
	Lajedo dos Eurípedes	01420.010042/2011-31	54160.002247/2013-80
	Casal I	01420.010048/2011-17	-
	Casal II	01420.010048/2011-17	-

	Gonzaga	01420.013986/2011-61	-
	Irecezinho	01420.010029/2011-82	54160.002690/2014-31
	Lageado II	01420.010062/2011-11	54160.002238/2013-99
	Lagedinho	01420.010063/2011-57	54160.002988/2014-62
	Lagedo do Pau d'Arco	01420.001039/2011-27	54160.002689/2014-15
	Largo	01420.010064/2011-00	-
	Macambira	01420.009786/2012-94	54160.002687/2014-18
	Patos	01420.010059/2011-99	-
	Salgada	01420.010038/2011-73	54160.002233/2013-66
	Volta Grande	01420.009788/2012-83	54160.002691/2014-86
<b>MULUNGU DO MORRO</b>	Alagadição	01420.006781/2014-71	-
	Alagadiço II	01420.006781/2014-71	-
	Baixa da Cainana	01420.006773/2014-25	-
	Caatinga do Egidio	01420.006772/2014-81	54160.000484/2017-23
	Caldeirão	01420.006785/2014-50	54160.000484/2017-24
	Cascavel	01420.006775/2014-14	54160.001304/2017-37
	Lagoa Damasceno	01420.006771/2014-36	-
	Lagoa Preta	01420.006783/2014-61	-
	Lagoa Vermelha	01420.006783/2014-61	54160.001305/2017-81
	Queimada da Onça	01420.006777/2014-11	54160.000484/2017-25
	Rosendo	01420.006785/2014-50	54160.000484/2017-24
	Umburaninha do Dionísio	01420.006779/2014-01	-
<b>P. DUTRA</b>	Coloudo Ramos	01420.000884/2010-02	-
<b>SÃO GABRIEL</b>	Boa Hora	01420.006552/2011-12	-
	Buqueirão dos Carlos	01420.006554/2011-01	54160.001254/2013-64
	Caroazal	01420.003823/2011-70	-
	Currálinho	01420.006558/2011-81	-
	Lagoinha	01420.003460/2006-13	-
<b>UIBAÍ</b>	Caldeirão	01420.002666/2006-18	-
	Lagoinha	01420.002662/2006-20	-
<b>XIQUE-XIQUE</b>	Ilha de Champrona	01420.101766/2019-41	-
	Vicentes	01420.001741/2006-23	54160.002941/2008-30

O *Quadro 8* aponta 148 Comunidades com certificado de reconhecimento enquanto Comunidade de Remanescentes de Quilombo – termo utilizado pela Fundação Cultural Palmares –, dessas, apenas 70 iniciaram a abertura do processo de regularização fundiária pelo INCRA. A disparidade numérica nos traz o questionamento em relação aos empecilhos ao avanço nos processos de titulação das terras. Faz-se importante uma breve discussão acerca dessa problemática, para que seja possível compreender os motivos que conduzem à estagnação do processo de reconhecimento dessas comunidades.

Sabe-se que, a partir da promulgação da Constituição Federal de 1988, bem como o surgimento posterior de outras normas constitucionais, a regularização territorial quilombola evoluiu do debate público para o cenário jurídico, quando ficou amplamente estabelecido que primeiramente as associações autointituladas enquanto remanescentes de quilombos deveriam requerer a expedição da certidão de autorreconhecimento à Fundação Cultural Palmares. O próximo passo para a política de reconhecimento quilombola fica a cargo do INCRA, através do processo de regularização fundiária, no entanto, como aponta o *Quadro 8*, os poucos processos abertos até o momento no TII permanecem sem avançar, não tendo nenhuma das terras tituladas pelo INCRA.

Importante ressaltarmos que, durante a década de 1980, no período de redemocratização, que culminava também com o centenário da abolição da escravidão, houve grande mobilização da comunidade negra no país, lutando por seus direitos à cidadania e pela normatização de instrumentos de combate ao racismo, o que resultou, em 1988, na criação, no Congresso Nacional, tendo a sanção do então presidente José Sarney, da Lei nº 7.668/88, que dá vida à Fundação Cultural Palmares, primeira instituição de Estado a tratar da questão racial, “para conduzir as complexas questões necessárias e mudar a realidade discriminatória e excludente da sociedade brasileira” (Araújo, 2008, p. 5).

No mesmo processo de lutas, houve a conquista de inserção, na Constituição de 1988, do reconhecimento do direito de posse às terras dos antigos quilombos por parte das comunidades negras. Houve então a importante fundamentação, nos Atos das Disposições Constitucionais Transitórias, da Constituição Federal de 1988, em seu art. 68, que versou sobre o direito desse povo à terra que já ocupa, devendo o Estado então emitir os devidos títulos. Após esse marco legal, os últimos anos foram de “reivindicações ao direito à permanência e ao reconhecimento legal da posse de suas terras ocupadas e cultivadas para moradia e sustento” dos afrodescendentes, chamados negros, em todo o território nacional, através de suas organizações em associações quilombolas (Leite, 2000; Monteiro; Garcia, 2010).

Em 1995, a Fundação Cultural Palmares, órgão ligado ao Ministério da Cultura, emitiu a portaria nº 25, visando estabelecer as normas para os trabalhos de identificação e delimitação das terras ocupadas por comunidades remanescentes de quilombos. Ademais, em novembro desse mesmo ano, o INCRA, autarquia do Ministério do Desenvolvimento Agrário, emitiu a portaria nº 307, em que elenca os meios para demarcar e expedir os devidos títulos às terras quilombolas inseridas em áreas públicas. A partir de então, inicia-se o estabelecimento dos critérios de certificação de autorreconhecimento das comunidades, e, através da Instrução Normativa nº. 57 do INCRA, torna-se possível conhecer os detalhes do processo de identificação, reconhecimento, delimitação, demarcação e titulação das terras ocupadas por remanescentes das comunidades quilombolas (termo utilizado na Instrução Normativa). Para tanto, disponibiliza-se no próprio *site* uma cartilha “Passo a passo da Titulação de Território Quilombola”, conforme descrito no *Quadro 9*.

**Quadro 9** – Passo a passo da Titulação de Território Quilombola

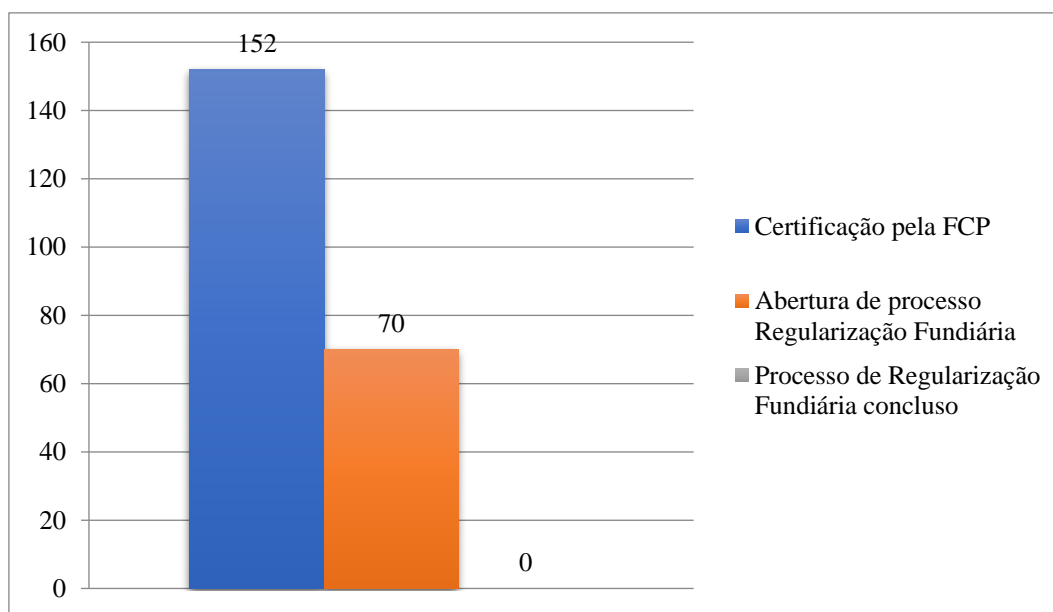
<b>ETAPAS</b>	<b>DESCRIÇÃO</b>
<b>AUTODEFINIÇÃO QUILOMBOLA</b>	A comunidade quilombola, assim como qualquer outro grupo social, tem direito à autodefinição. Para regularizar seu território, o grupo deve apresentar ao Incra a Certidão de Autorreconhecimento, emitida pela Fundação Cultural Palmares.
<b>ELABORAÇÃO DO RTID</b>	A primeira etapa de regularização fundiária quilombola consiste na elaboração do Relatório Técnico de Identificação e Delimitação (RTID), visando o levantamento de informações cartográficas, fundiárias, agronômicas, ecológicas, geográficas, socioeconômicas, históricas, etnográficas e antropológicas, obtidas em campo junto a instituições públicas e privadas. O RTID tem como objetivo identificar os limites das terras das comunidades remanescentes de quilombos.
<b>PUBLICAÇÃO DO RTID</b>	Os interessados terão o prazo de 90 dias, após a publicação e as notificações, para contestarem o RTID junto à Superintendência Regional do Incra, juntando as provas pertinentes. Do julgamento das contestações caberá recurso único ao Conselho Diretor do Incra Sede, no prazo de 30 dias a contar da notificação.
<b>PORTARIA DE RECONHECIMENTO</b>	A fase de identificação do território encerra-se com a publicação de portaria do Presidente do Incra que reconhece os limites do território quilombola no Diário Oficial da União e dos estados.
<b>DECRETO DE DESAPROPRIAÇÃO</b>	Nos casos em que há imóveis privados (títulos ou posses) incidentes no território, é necessária a publicação de Decreto Presidencial de desapropriação por Interesse Social. Os imóveis desapropriados serão vistoriados e avaliados conforme os preços de mercado, pagando-se sempre previamente e em dinheiro a terra nua, no caso dos títulos válidos, e as benfeitorias.
<b>TITULAÇÃO</b>	O Presidente do Incra realizará a titulação mediante a outorga de título coletivo, imprescritível e pró-indiviso à comunidade, em nome de sua associação legalmente constituída, sem nenhum ônus financeiro. É proibida a venda e penhora do território.

Fonte: Adaptado do *site* do Incra (publicação de 2020).

Assim, as questões de acesso à terra aos autointitulados remanescentes quilombolas passaram a ser mais evidentes no estado, quando, em 11 de outubro de 2013, é sancionada a lei que finalmente dispõe sobre a regularização fundiária das terras públicas estaduais, rurais e devolutas, que estejam ocupadas pelas CRQs, estabelecendo logo no primeiro artigo o reconhecimento da propriedade definitiva, sob título de domínio coletivo e pró-indiviso, expedido em nome da associação comunitária legalmente constituída, que represente a coletividade quilombola (art. 1º, caput e § 2º, da Lei do Estado da Bahia n.º 12.919, de 11 de outubro de 2013).

No entanto, mesmo com os avanços legais e normatizadores, no Território de Irecê existem atualmente 148 comunidades com certidões de autorreconhecimento expedidas, entretanto apenas 70 estão com processo de regularização fundiária aberto no INCRA, e nenhum título de domínio já expedido no território (atualizado em 16.05.2022, através do *site* do INCRA) – conferir *Gráfico 5*.

**Gráfico 5** – Quantitativo de Comunidades Certificadas e com Regularização Fundiária



Fonte: Elaborada pela autora.

Os dados expressos no *Gráfico 5* são um alerta em relação a possíveis entraves que existem na/para atuação do INCRA, no que tange ao cumprimento do que já é amplamente estabelecido em diversas disposições legais, fazendo-se de suma importância refletirmos sobre quais são e os porquês das dificuldades encontradas no processo, as quais levam à ineficácia da

devida titulação das terras quilombolas do território – um convite a pesquisas sociológicas, antropológicas.

#### 4.1.4 Comunidades do Território: dados sociodemográficos para uma breve caracterização identitária

Conhecer o panorama histórico das comunidades linguísticas, bem como suas características sociodemográficas, é imprescindível ao desenvolvimento da pesquisa de cunho sociolinguístico, visto que

muitas histórias têm de ser reconstruídas para que se venha a ter uma história do diversificado português brasileiro, inexoravelmente nascido do encontro da ‘linguagem adulterada de negros e índios’ e da Koiné portuguesa, além de outras línguas aqui chegadas que, necessariamente, por razões sócio-históricas e linguísticas, entrecruzaram-se [...]. (Mattos e Silva, 2004a, p. 12).

Com indica Mattos e Silva (2004), aqui, nesta seção, tentaremos apresentar, ainda de forma generalizada, dados sociodemográficos que favoreçam a compreensão do contexto atual dessas comunidades que ainda são rigorosamente marcadas pela segregação social, resistência e luta pela sobrevivência que atravessa a história dos negros e afro-descendentes no território brasileiro, em especial, no TII.

O Projeto ELiHS teve acesso a algumas comunidades e pôde vivenciar de perto o enfrentamento à falta de acesso a serviços básicos de saúde, educação, transporte, bem como uma base agrícola familiar presente nas Comunidades. A partir de então, passamos a entender aquelas Comunidades como espaço de vida e resistência dos seus membros que lutam para terem acesso e permanecerem na terra, e, ainda, é notável o campo como espaço de produção material e simbólica das condições de existência, de construção de identidades.

**Figura 14** – Retratos de membros das Comunidades em atividade de campo



Fonte: Acervo do Projeto ELiHS (divulgação autorizada).

Os retratos contidos na *Figura 14* marcam nitidamente a manutenção de tradições que seus antepassados trouxeram da África; no campo da agricultura, uma relação comunitária no uso da terra, refletindo um contexto de tradições próprias e histórias de múltiplas resistências, modos de vida próprios, condições econômicas de subsistência familiar/comunitária.

Mesmo inseridos numa sociedade compreendida enquanto contemporânea, as comunidades afro-brasileiras do TII representam relação social, política e simbólica que liga seus membros à sua terra, e conseqüentemente favorece a construção de sua identidade cultural, construindo um modo de vida particular e caracterizando o que Bortoni-Ricardo (2021) denomina de uma rede social mais isolada (*cf.* Figura 12). Assim, podemos afirmar que, apesar de nos autointitularmos enquanto sociedade moderna, engajada em um processo contínuo de transformação histórica, temos muitas comunidades, a citar as afro-brasileiras do TII, que se apresentam e constroem sua identidade dentro de um sistema conservador. Vejamos mais



alguns retratos que nos permitem conhecer essa realidade, muitas vezes paralela ao campo científico/universitário.

**Figura 15** – Casa do poço artesiano- Comunidade de Floresta I



Fonte: Acervo do Projeto ELiHS.

**Figura 16** – Chafariz de Abastecimento desativado – Comunidade de Segredo



Fonte: Acervo do Projeto ELiHS.

**Figura 17** – Cemitério – Comunidade de Cruzeiro



Fonte: Acervo do Projeto ELiHS.

**Figura 18** – Lajedo – Comunidade de Brejinho



Fonte: Acervo do Projeto ELiHS.

**Figura 19** – Exterior e interior da primeira escola da Comunidade de Floresta II



Fonte: Acervo do Projeto ELiHS.

**Figura 20**– Retratos de moradias das comunidades



Fonte: Acervo do Projeto ELiHS.

Os retratos das *Figuras 14 a 20* refletem um modo de vida simples. Esse perfil, em releituras, reflete um cenário de disparidades socioeconômicas/sociais que não representa apenas um simples legado do passado, haja vista ainda estar sendo perpetuado pela estrutura desigual de oportunidades sociais a que brancos e negros estão expostos no presente. As Comunidades do TII marcam a resistência do povo negro e aqui será caracterizada como afro-brasileira, descendentes de negros escravizados vindos de várias regiões do continente africano. Importante frisar que, em um contexto geral, comungando com o *continuum* proposto por Bortoni-Ricardo (2004, 2005, 2011),

[...] esses descendentes vivem principalmente no espaço rural brasileiro, *mas muitos núcleos estão incorporados às áreas periurbanas e urbanas do país*. Em função dessas diferenciações de localização espacial, essas comunidades caracterizam-se por apresentar níveis diferenciados de inserção e de contato com a sociedade. (Anjos, 1999, p. 35, grifo nosso).

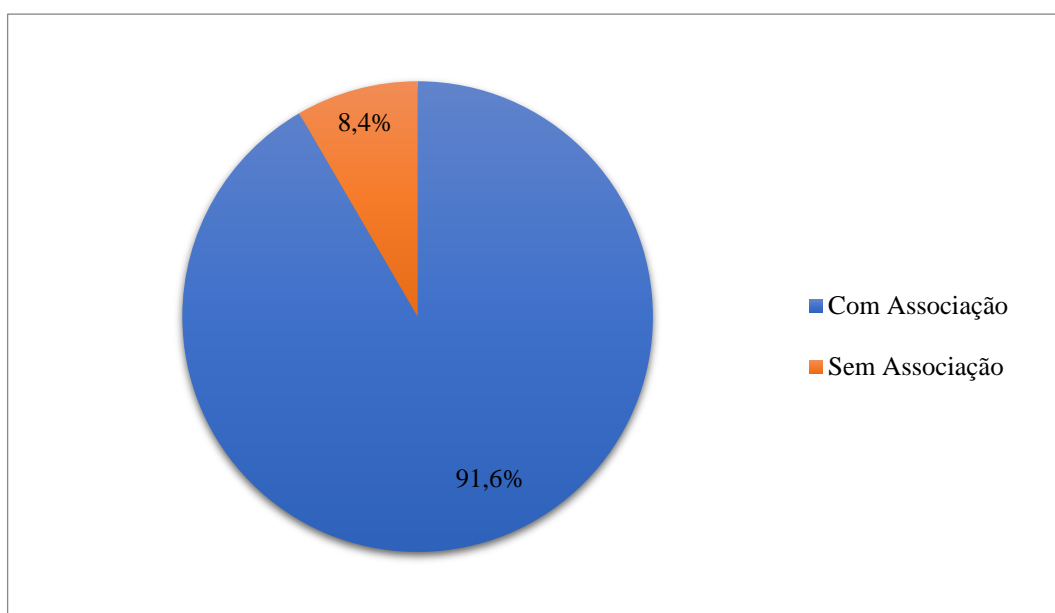
Dados levantados a partir do Projeto ELiHS apontam a existência de comunidades afro-brasileiras, localizadas majoritariamente em espaços rurais, em 19 municípios do Território de Irecê, não sendo computado apenas o município de Ipupiara. No site do Projeto, descrito na *seção 4.3*, será possível conhecer um pouco mais dos retratos das Comunidades, uma oportunidade singular para pessoas que, mesmo diante da distância geográfica e dificuldade de acesso, anseiam por conhecer mais realisticamente o atual cenário das comunidades afro-brasileiras, certificadas pela FCP enquanto remanescentes de quilombos. Certamente, através do recurso da fotografia, conseguiremos estabelecer uma comunicação multicultural, por meio de imagens, as quais permitirão captar e transmitir o que não é imediatamente transmissível no plano verbal, colaborando diretamente na construção simbólica dessas comunidades tradicionais; afinal, como afirma Jean Paul Sartre (1979), “na imagem, o objeto entrega-se em bloco e o olhar está certo disso – ao contrário do texto ou de outras percepções que me apresentam o objeto de uma forma frouxa, discutível e assim me levam a desconfiar daquilo que julgo ver”. Assim, acreditamos que esses retratos (expostos no *site* do Projeto) serão uma forma de imortalizarmos o registro de uma brava história de luta e resistência, que perdura até os dias atuais.

Aliado ao contexto visual, fazem-se necessários os dados estatísticos, que são fundamentais para o processo de caracterização sociodemográfica, complementando o objetivo do Projeto de interligar aspectos de natureza histórica e linguística. Nesse sentido, apresentaremos graficamente dados inéditos obtidos através de pesquisa, realizada

eletronicamente (via *Google Forms*), no âmbito do Projeto ELiHS, entre os anos de 2020 e 2021, em que foi possível coletar informações sociodemográficas em um conjunto de 109 comunidades das 148 certificadas pela Fundação Cultural Palmares, situadas no TII.

Iniciaremos com dados representativos do quantitativo de Associações, visto que a constituição legal da Associação é indispensável para o andamento do processo de regularização fundiária pelo INCRA, como já discutido na *seção 4.1.3*. A tabulação dos dados permitiu identificar que 91,6% das Comunidades estão organizadas em Associações<sup>42</sup>, compreendidas como uma forma de organização política e social que possibilita o diálogo e o trabalho coletivo, através de membros representativos. Os dados podem ser melhor observados no *Gráfico 6*.

**Gráfico 6** – Quantitativo de Associações nas Comunidades do TII



Fonte: Elaborada pela autora.

Os dados apresentados são apenas representativos, mas fica o alerta para as Comunidades que representam os 8,4% e que não sabem quais os processos protocolais para a criação dessas Associações: “A Comunidade não possui associação, a maioria dos moradores desejam ter orientações” (Representante de Sertão Bonito – João Dourado). Que vozes como essa, da Comunidade Sertão Bonito, ecoem nas esferas públicas, na tentativa de medidas

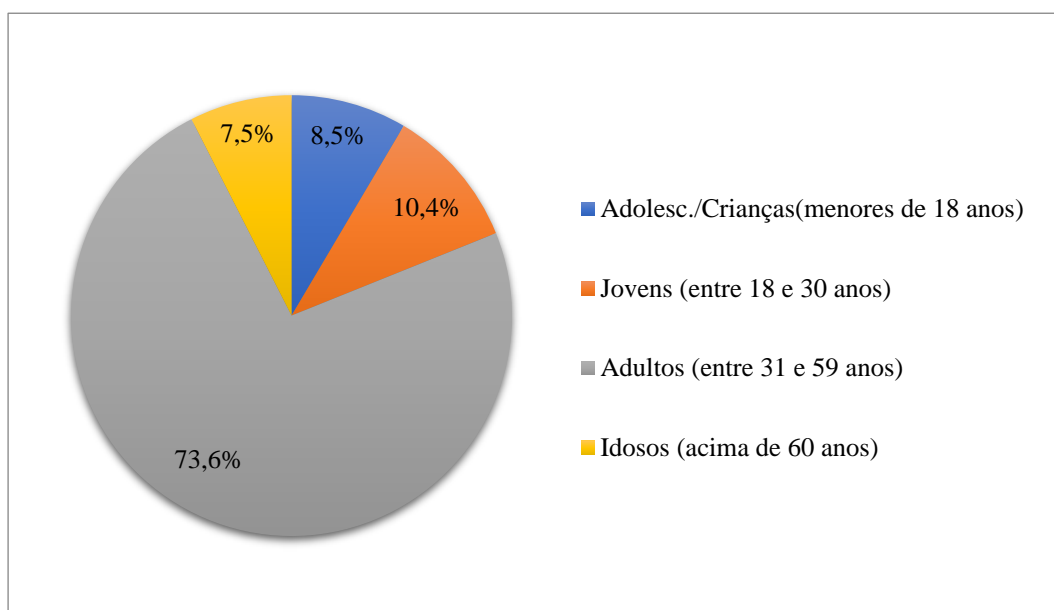
---

<sup>42</sup> A maioria das Associações das Comunidades não dispõem de espaço físico próprio, utilizando-se, muitas vezes, de espaços escolares ou outros espaços públicos.

interventivas, já que a representação através de Associações favorece um movimento mais forte e combativo que passa a pensar na valorização do sujeito, na cultura popular e na sua identidade enquanto comunidade tradicional afro-brasileira, além de uma organização que favorece a aquisição de políticas públicas e a promoção do respeito aos direitos humanos, tão urgentes nesse contexto de grupos que estão em condição de vulnerabilidade social.

Em relação ao número de habitantes, temos comunidades que registram 10 (dez) habitantes (a exemplo de Garapa, em América Dourada), e 02 (dois) mil habitantes (a exemplo de Prevenido, também em América Dourada), chegando a uma média geral de 321 habitantes/comunidade. A catalogação dos dados nos permite afirmar que seus membros compõem predominantemente a faixa etária de adultos (entre 31 e 59 anos), com 73,6%; seguidos dos jovens (entre 18 e 30 anos), com 10,4%; adolescentes e crianças, com 8,5%; e idosos, com apenas 7,5%, conforme *Gráfico 7*.

**Gráfico 7** – Quantitativo da proporção por grupos etários nas Comunidades do TII



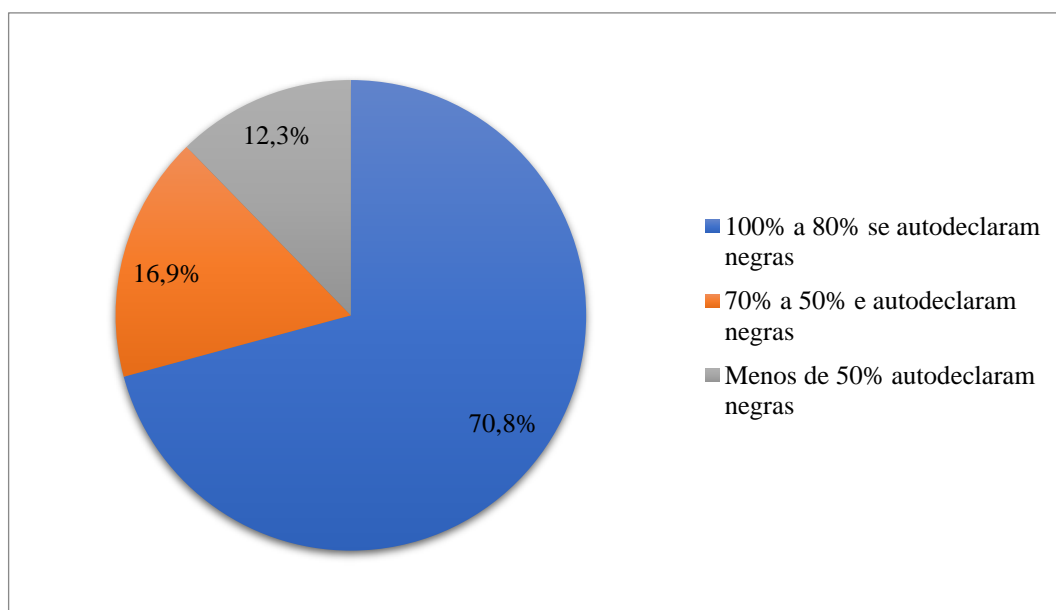
Fonte: Elaborada pela autora.

Entrecruzando os dados sociais com as pretensões linguísticas deste Projeto, abrimos o parêntese para o fato de termos apenas 7,5% dos membros numa faixa etária a partir de 60 anos, sendo que a idade é uma das super categorias sociais – ou variáveis previsoras socioculturais, como trataremos nas análises, *seção 5.3* – para os estudos sociolinguísticos, e suas implicações nos estudos quantitativos correlacionam-se diretamente com o processo de

mudança linguística e, intuitivamente, com a preservação de traços indicadores de possíveis contatos linguísticos.

Um outro dado estatístico que trazemos aqui refere-se à autoafirmação dos membros enquanto negros e/ou afrodescendentes, tendo no TII um surpreendente quantitativo de comunidades que se autodeclaram enquanto remanescentes de quilombolas, mas que não se autodeclaram enquanto afrodescendentes (*cf.* Gráfico 8).

**Gráfico 8** – Quantitativo das Comunidades que se autodeclaram enquanto negro e/ou afrodescendentes no TII

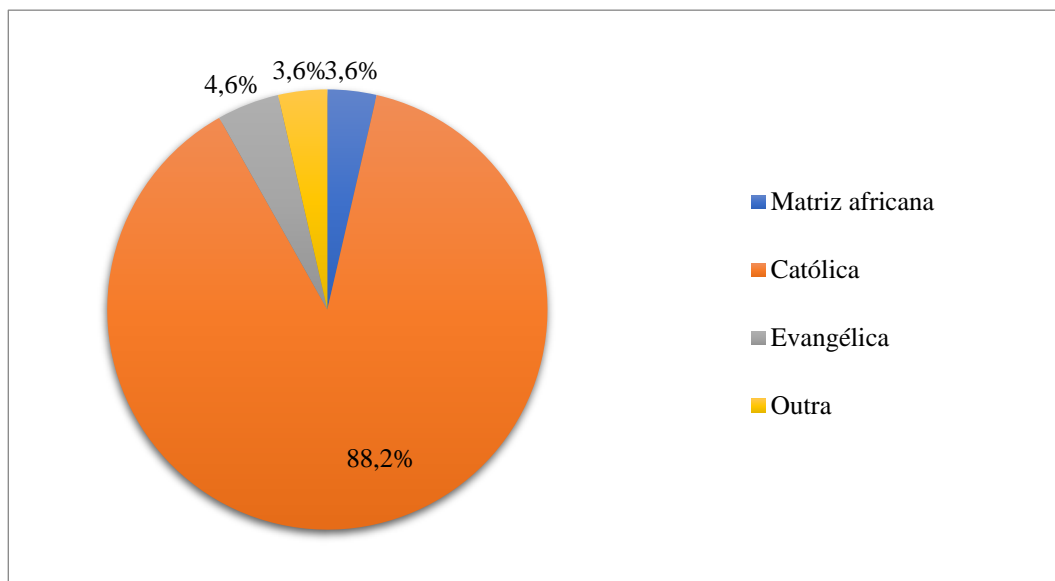


Fonte: Elaborada pela autora.

Os números apresentados contribuem para uma reflexão em relação à forte (re)produção dos estereótipos que nascem e resultam de/em fortes práticas de exclusão às pessoas negras, fruto de uma origem histórica que beneficia grupos esbranquiçados e por meio dele constrói um suposto poder hegemônico com relação ao restante da sociedade, instrumentalizando o racismo por meio do imaginário social, em uma sociedade multirracial e mestiçada.

Ainda em um contexto de autorreconhecimento da identidade cultural de seus povos – quilombolas –, temos dados que apontam que 88,2% comungam com os princípios da igreja católica; 3,7%, com princípios de religiões de matriz africana (Candomblé e Umbanda); 4,6%, com igreja evangélica; e 3,6%, com outras religiões.

**Gráfico 9** – Quantitativo das religiões predominantes nas Comunidades do TII



Fonte: Elaborada pela autora.

**Figura 21** – Retratos das religiões predominantes nas Comunidades do TII



Fonte: Acervo do Projeto ELiHS.



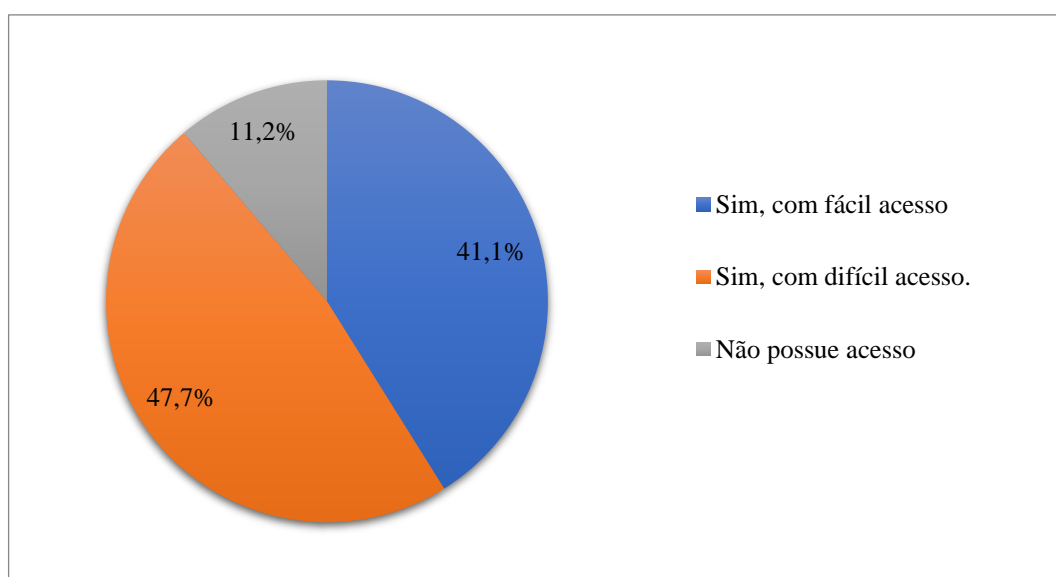
No que tange à religiosidade, as Comunidades são marcadas, em sua maioria, por influências históricas e culturais de matrizes católica e protestantes, visto que, no presente, os costumes da cultura religiosa de matriz africana não são claramente manifestados, envolvendo questões de preconceito e intolerância religiosa. Aproveitamos o espaço para destacar as práticas religiosas sincréticas de cura, como os benzimentos, que são realizados por meio de ritos católicos, mesclados com saberes da cultura afro-indígena. Como afirma João Reis (1996), essa disponibilidade para mesclar culturas era um imperativo de sobrevivência, também refletida na habilidade demonstrada pelos quilombolas de constituir alianças sociais dentro e fora do grupo, o que talvez justifique o quantitativo direcionado às igrejas católicas e protestantes.

Os dados apresentados nos *Gráficos 8 e 9* e os retratos apresentados na *Figura 21* fornecem subsídios para importantes discussões em torno do conceito de identidade, atrelado aos modelos ditados pela sociedade contemporânea, visto que Hall (2003, p. 13) afirma que

o sujeito assume identidades diferentes em diferentes momentos, identidades que não são unificadas ao redor de um 'eu' coerente. Dentro de nós há identidades contraditórias, empurrando em diferentes direções, de tal modo que nossas identificações estão sendo continuamente deslocadas

Não consideramos as Comunidades do TII como isoladas, por razões já explicadas anteriormente, no entanto os dados apresentados nos *Gráficos 9 e 10* permitem caracterizar essas Comunidades com [-] informatizadas, visto que 11,2% não possuem acesso telefônico, somando às 41,1% que o possuem com difícil acesso.

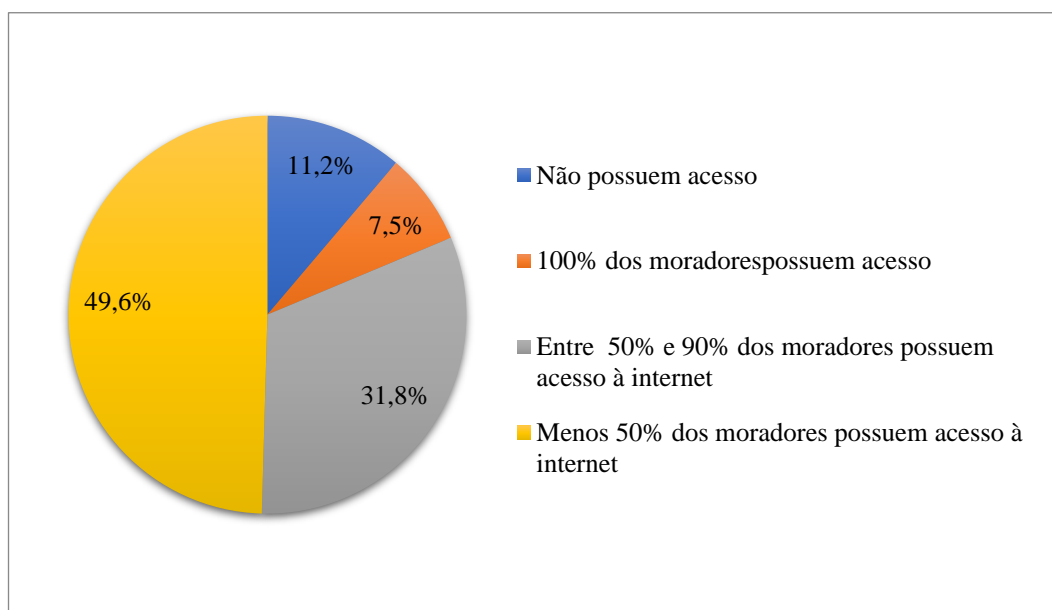
**Gráfico 10** – Acesso telefônico nas Comunidades do TII



Fonte: Elaborada pela autora.

Em relação ao acesso à internet, temos 11,2% das Comunidades sem nenhum tipo de acesso, somando às 31,8% com registro de que apenas menos de 50% da população/membros possuem acesso<sup>43</sup>, evidenciando quantitativamente um uso ainda restrito da internet, o que configura a exclusão desse grupo étnico em ambientes digitais. Em contexto de pandemia do novo Coronavírus, também se questiona o acesso desses adolescentes a espaços escolares, os quais, desde março de 2020, têm trabalhado de forma remota – mais um convite a pesquisas.

**Gráfico 11** – Acesso à internet nas Comunidades do TII



Fonte: Elaborada pela autora.

A ausência da escola no campo é um outro dado que merece destaque, visto que tem se configurado como um fator que favorece o êxodo rural e consequente aumento da taxa de urbanização, bem como a falta de perspectiva para o viver no e do campo.

<sup>43</sup> Muitas das Comunidades, por estarem distantes de suas sedes, possuem apenas a opção da internet via rádio, ainda de alto custo ao se comparar à renda dessas famílias, e apresentando muita instabilidade na conexão à rede.

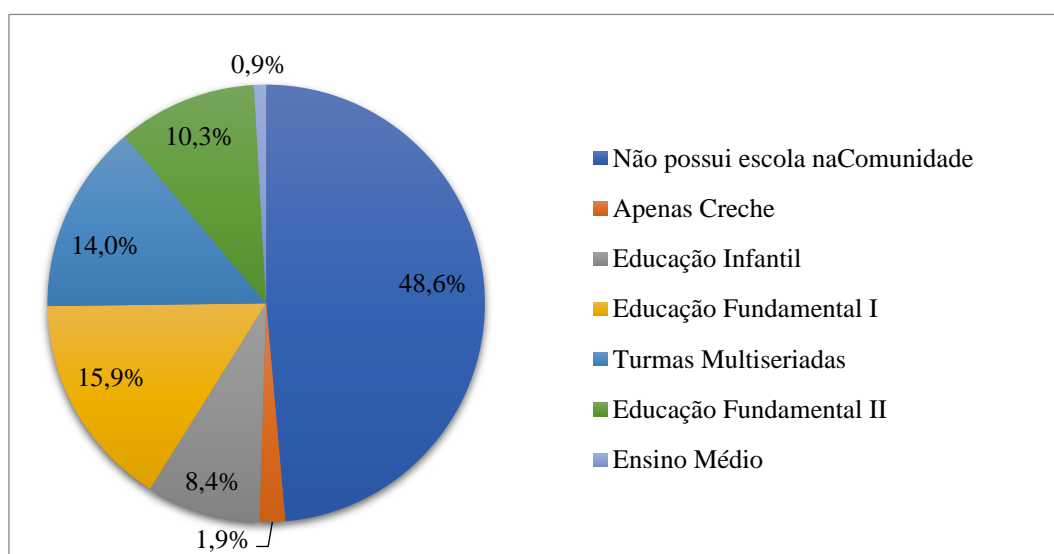
**Figura 22** – Escola Ensino Fundamental I desativada – Comunidade de Cruzeiro



Fonte: Acervo do Projeto ELiHS.

No TII, temos, surpreendentemente, 48,6% das Comunidades sem escolas ou com escolas desativadas, como na *Figura 21*. O quantitativo de escolas em funcionamento, em sua maioria apenas com Ensino Fundamental I, conforme *Gráfico 12*, favorece o baixo índice de escolaridade de seus membros, apontado no *Gráfico 13*.

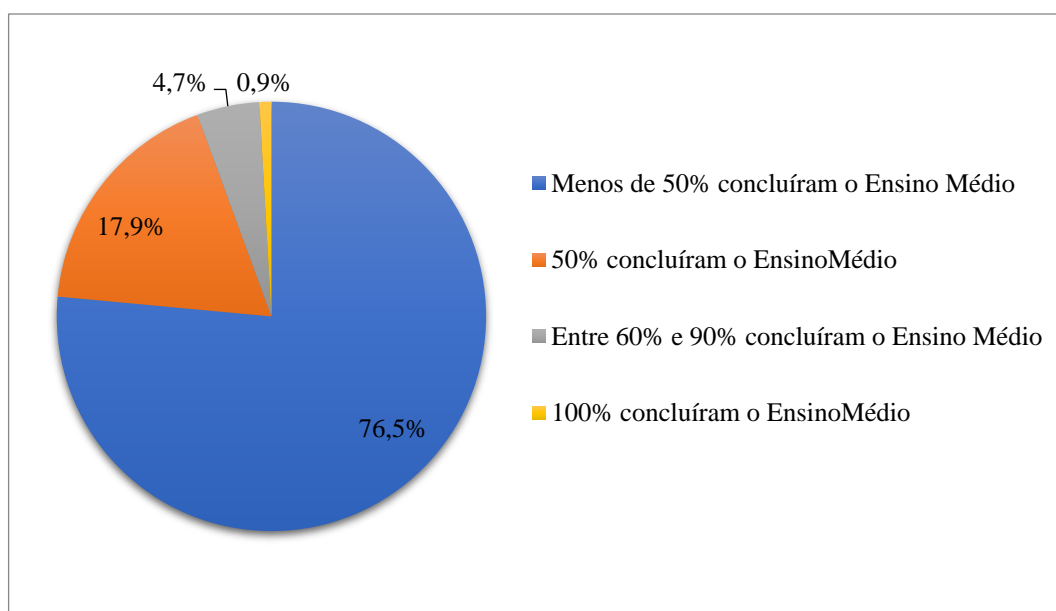
**Gráfico 12** – Níveis de ensino nas Comunidades do TII



Fonte: Elaborada pela autora.

As Comunidades que compõem o TII têm escolas que atendem no máximo o Ensino Fundamental II (10,3%); normalmente o Ensino Médio (0,9%) acontece como extensão do colégio que se encontra na sede do município. É um dos dados que revela a concentração da vida social na sede do município e/ou o baixo nível de escolarização, como uma matriz que fortalece a lógica do inchaço urbano, do êxodo rural e da falta de perspectiva para o viver no e do campo. Os dados apontam que o panorama histórico da educação escolar brasileira na sua vertente rural nos deixou como herança um quadro de precariedade no funcionamento da escola no/do campo, tanto em relação à infraestrutura e aos espaços físicos inadequados, quanto à má distribuição das escolas no espaço geográfico.

**Gráfico 13** – Quantitativo de pessoas/membros que concluíram o Ensino Médio nas Comunidades do TII



Fonte: Elaborada pela autora.

Os dados apontam que apenas 0,9% das Comunidades registram que 100% de seus membros concluíram o Ensino Médio. Em contrapartida, 76,5% registam que menos de 50% de seus membros concluíram o Ensino Médio. Como relatado anteriormente, a ausência de escolas e o fato de o Ensino Médio, em sua grande maioria, ser uma extensão dos colégios que se encontram na sede do município favorecem e explicam os baixos índices de escolaridade.

Diante dos quantitativos apresentados, é possível, ainda preliminarmente, visualizar um desenho social das Comunidades do TII, que, mesmo diante das prerrogativas normativas, ainda são invisibilizadas pelo poder público local, que não possui registro algum sobre esse grupo étnico tão representativo em sua constituição territorial. Portanto, os inéditos e ainda

incipientes dados quantitativos aqui apresentados são um convite a pesquisas, das mais diversas perspectivas teóricas, para que seja possível colaborarmos para constituição de um retrato não apenas sociolinguístico, mas também social, cultural e histórico, dando visibilidade a contextos, como este apresentado, que esteja “para além do litoral”, numa perspectivava de interiorização, já que, como se afirma antecipadamente na epígrafe, “[...] a minha aldeia é tão grande como outra terra qualquer [...]” (Pessoa, 1993 [1946]).

#### **4.2 Projeto Estudos Linguísticos e Históricos do Sertão (ELiHS)**

Como já afirmam Santana, Araújo e Freitag (2018a), quando envoltos em um cenário de padronização e normatização das variedades linguísticas brasileiras, nós, sociolinguistas, somos tomados por uma latente necessidade de discutirmos sobre a diversidade linguística brasileira, em especial as variedades rurais, ainda afastadas dos grandes centros urbanos, na tentativa de evitar que as principais características do vernáculo brasileiro popular sejam apagadas em meio à crescente urbanização e aos avanços tecnológicos.

Nesse contexto de anseios e no afã de proporcionar material oral para análise de descrições, especialmente no campo da linguística, delineia-se o Projeto *Estudos Linguísticos e Históricos do Sertão* (ELiHS), coordenado pela professora Dayane Moreira Lemos<sup>44</sup>, no Departamento de Ciências Humanas e Tecnologias (DCHT) da Universidade do Estado da Bahia (UNEB), *campus* XVI – Irecê/BA.

O projeto tem por objetivo basilar contribuir para conhecimento da realidade linguística brasileira e, de forma específica, sobre o português popular falado em comunidades rurais afro-brasileiras do semiárido baiano, especificamente as situadas no Território de Identidade de Irecê. Tal objetivo surge em concatenação aos princípios da teoria do contato linguístico, discutida na *seção* 2.3.2, que nos permite afirmar que as variedades orais do português popular rural oferecem campos de trabalhos bastante promissores para o conhecimento da configuração sócio-histórica do Português Brasileiro, dado o fato de esse ter sido gestado sob o efeito de amplo contato linguístico, durante o período Colonial, e ter sido

---

<sup>44</sup> Atua como professora substituta no Colegiado de Letras, desde 2017, com contrato renovado até 2024. Ciente da sua passagem temporária, cabe frisar que o Projeto ELiHS, mesmo após a finalização contratual da coordenadora, permanecerá sediado no Departamento, sob a coordenação do professor Dr. Joabson Figueiredo, atual coordenador do Programa de Pós-graduação em Estudos Africanos, Povos Indígenas e Cultura Negra (PPGEAFIN), e as subcoordenações, hoje organizadas em três frentes, a saber: de atividades em campo; de Transcrições, Revisões e Armazenamento; e de Comunicação e Audiovisual.

alvo também de um contato mais recente com variedades padrões do PB, através dos processos de migrações (Lucchesi, 2009, 2003; Mattos e Silva, 2008a, 2001; entre outros).

Mattos e Silva (2008a, p. 23) nos apresenta indícios, já apresentados nesta tese, de que o principal elemento difusor do português no Brasil seria a 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. Portanto, a autora acredita que a reconstrução do passado do PB não pode deixar de estar atrelada ao conhecimento detalhado dos variados aspectos da história social nos espaços brasileiros e dos avanços, na atualidade, das teorias de contato linguístico. Ainda sob a perspectiva teórica da renomada Rosa Virgínia Mattos e Silva, acreditamos que, numa tentativa de compreensão em relação à reconstituição do português popular brasileiro,

[...] a busca de seu percurso histórico tem de ser feita não fundada em *corpora* escritos, organizável *ad hoc*, como para o *português culto brasileiro*, como é óbvio, mas num processo de reconstrução do tipo – que designarei metaforicamente – arqueológico, em que, de evidências dispersas, calçadas pelas teorias sobre o contacto linguístico e pela história social do Brasil, se possa chegar a formulações convincentes. Percurso análogo, *mutatis mutandis e modus in rebus*, ao da reconstrução do chamado ‘latim vulgar’, cuja principal fonte de estudo é o seu resultado, as línguas românicas (Mattos e Silva, 2008a, p. 24, grifos da autora).

Entende-se, dessa forma, como uma das perspectivas possíveis para a tentativa reconstrução desse português popular, a de buscar essas origens, do presente para o passado, segundo sugere Mattos e Silva (2008a), que aponta como um caminho o estudo feito com base em gravações de fala de comunidades orais urbanas e rurais, com histórias linguísticas diversificadas: “para a reconstrução do passado do português popular brasileiro, pesquisar no espaço brasileiro as variedades conviventes hoje *sobretudo as dos não escolarizados das diversificadas áreas rurais do Brasil*” (Mattos e Silva, 2008a, p. 29, grifo nosso).

Essas reflexões foram trazidas em uma comunicação, intitulada *Para a história do português culto e popular brasileiro: sugestões para uma pauta de pesquisa*, apresentada ao III Seminário do Projeto Para a História do Português Brasileiro (PHPB). Nesse contexto, Mattos e Silva (2008a, p. 28-29) sugere nove pautas de pesquisas, dentre as quais destacamos:

[...]  
e. para a reconstrução do passado do *português popular brasileiro*, pesquisar no espaço brasileiro as variedades conviventes hoje sobretudo as dos não-escolarizados das diversificadas áreas rurais do Brasil;  
[...]

g. enquadrar a formação do português brasileiro no contexto das teorias do contacto lingüístico, para ultrapassar a perspectiva antiga de busca das ‘influências’ africanas e indígenas;

[...]

i. reconstruir, detalhadamente e com a precisão possível, a sóciohistória lingüística das diversas áreas brasileiras cobertas pelo Projeto, considerando as línguas que aí estiveram em contacto, os movimentos demográficos, a ausência/presença da escolarização e, conseqüentemente da escrita, como elemento normativizador (Mattos e Silva, 2008a, p. 29, grifos do autor).

Ainda nesse contexto, não podemos deixar de relatar que o Projeto também se inspira e se fundamenta em duas das quatro vias propostas por Houaiss (1985) para a investigação da língua portuguesa no Brasil, a saber:

- i. **A via 2**, que consiste no “mapeamento confiável da dialectologia brasileira a exemplo do que sonhou Antenor Nascentes e realizou parcialmente Nelson Rossi com seus colaboradores – mapeamento do qual, pelas igualdades unitárias e globalizantes, será possível ‘recapitular’ o processo passado que terá gerado o presente descrito por essa dialectologia”; e
- ii. **A via 3**, que propõe o “incremento da dialectologia vertical em tantos quanto possíveis grandes centros urbanos e focos rurais antigos, a fim de se poder ver a interinfluência entre o rural e o urbano na transmissão adquirida e induzida”.

Nessa perspectiva, o Projeto se justifica, considerando-se que estudiosos de diversas perspectivas teóricas (Mattos e Silva, 2001; Houaiss, 1985) têm ressaltado a importância de buscar na socio-história respostas mais precisas que possam explicar a singularidade da nossa gramática no âmbito das línguas românicas. Assim, o Projeto ELiHS busca atender uma importante vertente de pesquisa para a recuperação da fragmentada história do português popular brasileiro no Sertão Baiano, do presente para o passado, somando esforços a outros projetos pioneiros – como o *Corpus* Eletrônico de Documentos Históricos do Sertão (CEDOHS/UEFS)<sup>45</sup>, Projeto Vertentes (UFBA)<sup>46</sup>, na tentativa de um estudo vertical das variantes populares do PB, não só as urbanas.

<sup>45</sup> Conferir site: <http://www5.uefs.br/cedohs/view/home.html>

<sup>46</sup> Conferir site: <https://vertentes.ufba.br/>

Assim, o Projeto busca contribuir, assim como tantos outros projetos<sup>47</sup>, com o processo de reconstrução e descrição linguística do PB, tentando conciliar aspectos de natureza sócio-histórica com aspectos de natureza linguística, já que

como sabemos, o percurso para a reconstrução do passado do *português brasileiro popular* não será o mesmo utilizável para a reconstrução do passado do *português brasileiro culto*, que se esteia numa tradição escrita. O *português popular brasileiro* fez-se e faz-se, ainda, não tanto quanto antes, é claro, na oralidade. (Mattos e Silva, 2008, p. 23, grifos da autora).

Baseando-nos nesses princípios norteadores, entendemos, no âmbito do Projeto ELiHS, a importância do processo de compreensão da socio-história linguística brasileira e dos estudos direcionados ao multilinguismo, na difusão da língua portuguesa no país, considerando cada área geográfica. Sendo assim, como afirma Mattos e Silva (2008, p. 30), “[...] que Deus seja louvado; porque o trabalho, a meu ver, será muito, de muitos e não será por pouco tempo”, e, dentre esses muitos, queremos apenas ser mais 01 (um) a somar, a acreditar na expansão da compreensão da socio-história do Português Popular do Brasil, a partir da constituição de um amplo registro do PAB do interior, para além do litoral, numa perspectiva sincrônica, do presente para o passado.

#### 4.2.1 Frentes de trabalho

Com o intuito de traçar um panorama sociolinguístico da fala popular de comunidades rurais afro-brasileiras situadas no Território de Identidade de Irecê – Ba, o Projeto ELiHS traz à tona a importância do mapeamento e da documentação do português falado nessas comunidades, variedade linguística que ainda não foi alvo de estudos, conferindo ao Projeto um grau de ineditismo. Assim sendo, o ELiHS vislumbra, a longo prazo, as seguintes frentes de trabalho, conforme descrito no *Quadro 10*.

---

<sup>47</sup> Aqui citamos o Projeto *O português falado em comunidades rurais afro-brasileiras no recôncavo baiano* (em execução), coordenado pela professora Dr.<sup>a</sup> Silvana Araújo, sediado na Universidade Estadual de Feira de Santana (UEFS), tendo como objetivo geral documentar e descrever comunidades de falas e de práticas (nova perspectiva no campo linguístico) afro-brasileiras da região do Recôncavo da Bahia, por meio de registros fônicos e de vídeos, realizando e fomentando pesquisas que tratem da participação de línguas e culturas africanas na formação do português brasileiro (PB).



Quadro 10 – Frentes de trabalho do Projeto ELiHS

FRENTES DE TRABALHO		DESCRIÇÃO	CRONOGRAMA DE EXECUÇÃO
MAPEAMENTO DAS COMUNIDADE AFRO-BRASILEIRAS DO TII	FASE 1	Mapeamento das comunidades afro-brasileiras presentes no Território de Irecê, certificadas pela Fundação Cultural Palmares (FCP) enquanto remanescentes de quilombos	2022
	FASE 2	Descrição sócio-histórica das comunidades afro-brasileiras do Território de Irecê	2022-2023
	FASE 3	Registo fotográfico representativo das comunidades afro-brasileiras do Território de Irecê.	2023-2024
COMPOSIÇÃO DE CORPUS ORAL	FASE 1	Formação de banco <i>online</i> , com amostras representativas de fala de residentes, <b>não escolarizados</b> , de comunidades afro-brasileiras do Território de Identidade de Irecê.	2020-2024
	FASE 2	Formação de banco <i>online</i> , com amostras representativas de fala de residentes, <b>escolarizados</b> , de comunidades afro-brasileiras do Território de Identidade de Irecê.	2025-2027
	FASE 3	Etiquetagem morfológica e anotação sintática do <i>corpus</i> <sup>48</sup>	A partir de 2027
O ELiHS NA SALA DE AULA	FASE 1	Criação de estratégias pedagógicas para o ensino de Língua Portuguesa a partir dos resultados estatísticos de <i>corpus</i> oral das comunidades do Projeto.	2023-2027
	FASE 2	Desenvolvimento de jogos interativos <i>online</i> enquanto proposta pedagógica para o Ensino de Língua Portuguesa, numa efetiva aproximação entre realidade e virtualidade, no que tange às comunidades do Território.	2023-2027
	FASE 3	Desenvolvimento e aplicação de curso extensionista para uso didático do jogo <i>online</i> , em escolas quilombolas do Território.	2024-2027

Fonte: Elaborado pela autora.

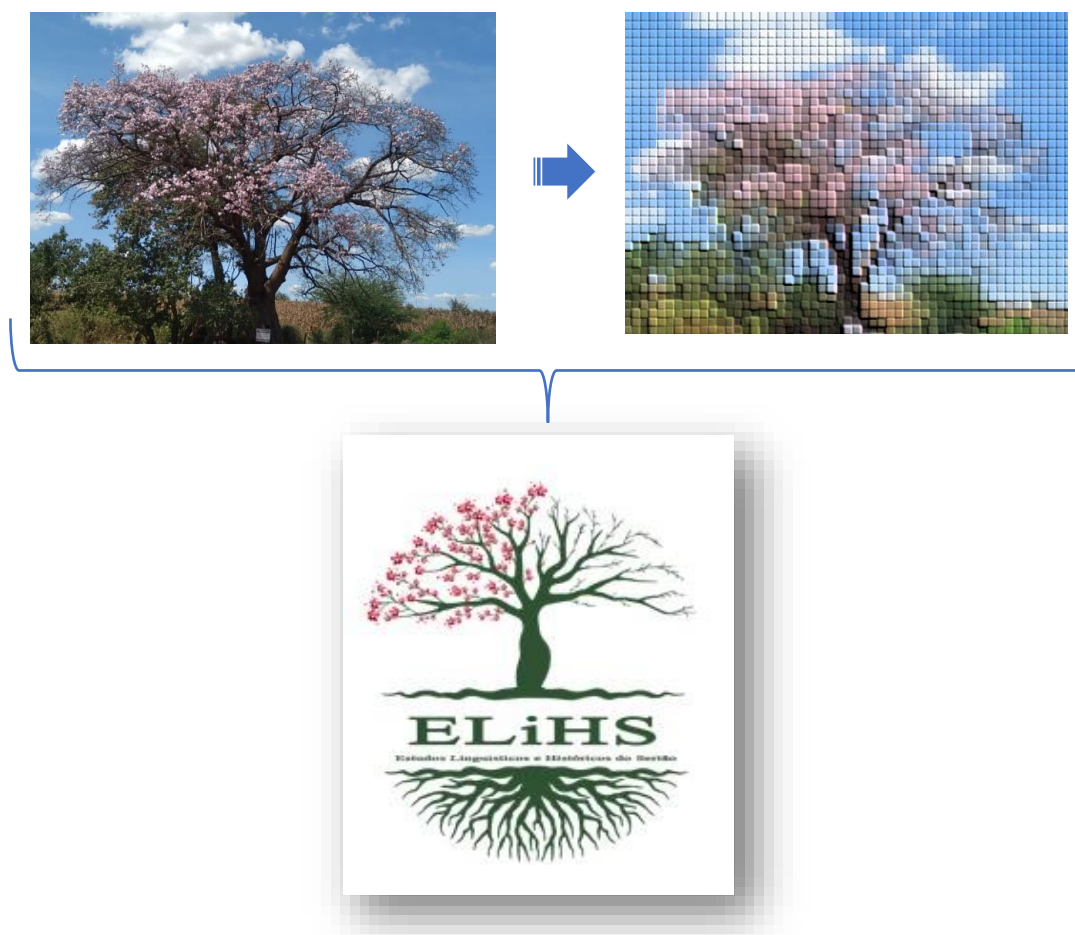
Essa base empírica permitirá ao Projeto ELiHS constituir um perfil sociolinguístico da fala popular de comunidades afro-brasileiras do Território, contribuindo, com isso, para o avanço do conhecimento referente a seus processos de variação e das possíveis tendências de mudança no campo linguístico, caracterizando o caráter multicultural e pluriétnico do português popular do Brasil, evitando, assim, que as principais características do vernáculo brasileiro

<sup>48</sup> Proposta de organização de um *corpus* anotado morfológica e sintaticamente para a execução de buscas automáticas e/ou semiautomáticas, bem como a recuperação rápida e confiável de grande quantidade de dados para pesquisas linguísticas. Seguiremos o padrão adotado no âmbito do *Corpus* Anotado do Português Histórico Tycho Brahe ([tycho.iel.unicamp.br](http://tycho.iel.unicamp.br)), coordenado por Charlotte Galves (UNICAMP). Para maiores esclarecimentos sobre anotação sintática, sugerimos a leitura de *Syntactic Annotation System*. Disponível em: <http://www.tycho.iel.unicamp.br/~tycho/corpus/manual/syn-frm.html>.

popular se apaguem, antes mesmo de serem documentadas. Além disso, a proposta do *Projeto ELiSH na sala de aula* traz uma inovação no que tange a interrelação entre a Sociolinguística Quantitativa e a Sociolinguística Educacional, em uma proposta de compromisso com a Educação, através de uma aplicação didático-pedagógica do resultado de pesquisas a serem realizadas no âmbito do Projeto, entendendo, como Ilari (2003, p. 111), que um dos principais desafios da Linguística no século XXI é “mudar a cara do ensino da língua materna”.

Portanto, a partir da defesa pública da tese, o Projeto tornar-se-á público e representado através da imagem da Barriguda (*Ceiba glaziovii*), árvore simbólica do Sertão baiano, conforme *Figura 23*.

**Figura 23**– Logomarca do Projeto ELiHS: do real à criação



Fonte: Acervo do Projeto ELiHS.

A *Figura 23* caracteriza-se como a inspiração para criação da logomarca do Projeto. Essa árvore está localizada na estrada da cidade de São Gabriel, sede de 05 (cinco) comunidades afro-brasileiras, e muitas outras há em vários pontos do território, sendo uma marca de força e

resistência para o povo sertanejo. Segundo Lemes (2011, p. 3, grifo nosso), a *Ceiba speciosa* St. Hil., pertencente à família *Malvaceae*, mais conhecida como paineira-rosa, paineira, árvore-de-paina, barriguda, dentre outras denominações, “é uma espécie nativa em matas secas no Brasil Central, que apresenta rápido crescimento e floração espetacular. Seu tronco abaulado e com acúleos, é característica marcante da espécie que também *se tornou bastante conhecida como barriguda*”. Sendo assim,

*a barriguda é uma espécie típica da Caatinga [...]. Trata-se de uma árvore de grande porte, que pelo formato dilatado de seu tronco na base formando uma barriga, assemelha-se ao baobá. O caule, assim como nos cactos, armazena água, uma adaptação para a planta viver em ambientes com forte estresse hídrico, como a Caatinga [...]. As flores estão reunidas em inflorescências terminais e são brancos-rosadas, possuindo numerosos e pronunciados estames [...]. É uma árvore majestosa [...]* (Stehmann; Faria; Bragioni, 2019, p. 32, grifo nosso).

A Barriguda é uma árvore representativa em toda a caatinga por sua força, imponência e resistência diante de longos períodos de seca. Sua “barriga” deposita uma quantidade razoável de água que garante sua sobrevivência em meio aos tempos difíceis de seca, permitindo o seu florescimento, que marca a força da resistência do sertanejo. Assim como a Barriguda, o Projeto ELiHS pretende resistir diante das dificuldades em se desenvolver pesquisa no Brasil, apresentando, ainda que timidamente, suas flores e seus avanços, abrindo campos promissores para desenvolvimento de pesquisas, com uma proposta norteadora de interiorização da descrição linguística. Sendo assim, para cada etapa do Projeto, teremos uma imagem representativa, vejamos:

**Figura 24** – Logomarca do Projeto ELiHS: fases de constituição



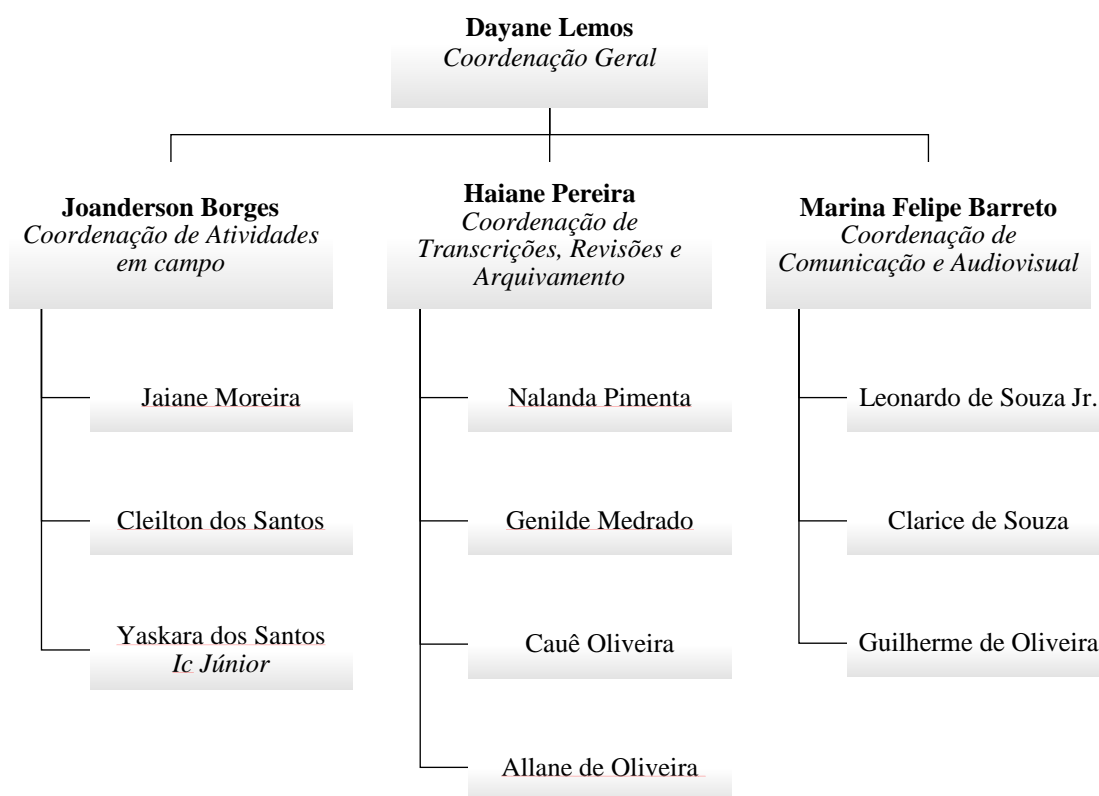
Fonte: Acervo do Projeto.

As flores apresentadas na *Figura 24* representam, ilustrativamente e progressivamente, o caminho a ser percorrido, bem como a abertura de possibilidades de estudo que favorecerão

o avanço do Projeto em suas fases, quando se dará o preenchimento das flores em seus caules, representando a conclusão do objetivo central, em suas fases 1, 2 e 3 (cf. *Quadro 10*).

O projeto conta atualmente com uma equipe composta por 01 (uma) coordenadora e 13 (treze) participantes, os quais são organizados em subcoordenações, a fim de uma maior organização das demandas de trabalho, bem como treinamentos (gravações e transcrições). As atribuições são direcionadas por área de interesse, que são definidas nas reuniões que acontecem quinzenalmente, nas sextas-feiras, momento em que realizamos grupos de discussão, montamos agendas, dialogamos sobre avanços e novas demandas, bem como possíveis dificuldades.

**Figura 25** – Organograma da equipe ELiHS: coordenação e subcoordenações



Fonte: Elaborada pela autora.

A coordenação de *Atividade de campo* se responsabiliza pelo treinamento e acompanhamento das gravações realizadas em campo, nas comunidades demarcadas pela equipe enquanto prioritárias (critérios a serem descritos na seção 4.2.2.3); a de *Transcrição, revisão e arquivamento* se responsabiliza pelo treinamento, organização das gravações realizadas em campo e encaminhamento para o processo de transcrição (critérios a serem descritos na seção 4.2.2.4), por fim, arquivamento no *drive* do Projeto; a de *Comunicação e Audiovisual* se

responsabiliza pela transferência do acervo, posto no drive, para plataforma *online*, bem como sua atualização, além de se responsabilizar, juntamente com a coordenadora geral, pelas estratégias de criação do espaço interativo, voltado ao ensino de língua portuguesa, iniciando, atualmente, com o jogo *online* para ensino da concordância nominal de número, fenômeno a ser descrito e analisado quantitativamente nesta tese.

A equipe que inicia com 02 (dois) bolsistas e hoje é apresentada por um quantitativo de 13 participantes, entre bolsistas e voluntários, é resultado de intenso trabalho que envolve não apenas a divulgação do Projeto para o Território, mas também o alcance de Editais que permitem o financiamento de bolsas de monitorias. Vejamos um pouco do reflexo do nosso trabalho:

(i) **Edital 15/2023** – Projeto de extensão

O Processo Seletivo, cujas inscrições estão abertas pelo presente, visa a concessão de Bolsas de Iniciação à Extensão a estudantes devidamente matriculadas/os nos cursos de graduação da UNEB, para participação em Projetos de Extensão Universitária. A seleção de bolsistas de iniciação à extensão visa propiciar aos estudantes a formação integral e cidadã a partir do envolvimento em ações de extensão voltadas à comunidade externa. Neste Edital, obtivemos a **aprovação de 02 (duas) bolsas de monitorias**, Clarice de Souza e Guilherme de Oliveira, que desenvolverão atividades remuneradas até dezembro/2023.

(ii) **Edital 73/2022** – Projeto de extensão

Edital referente à abertura das inscrições para apresentação de propostas para obtenção de bolsas de pesquisa e extensão no âmbito do Programa AFIRMATIVA de Pesquisa e Extensão nos Campi da UNEB. O Programa AFIRMATIVA, conformado às políticas de pesquisa e extensão da UNEB, ao inserir estudantes que ingressaram na UNEB através da reserva de vagas para negros(as) e sobre vagas para indígenas; quilombolas [...] no desenvolvimento inicial, orientado e supervisionado de atividades de elaboração e execução de projetos de pesquisa, extensão e de difusão do conhecimento, tem por objetivo possibilitar a esses estudantes uma forma específica de associação e ambientação coletiva que dê suporte material, intelectual e subjetivo ao desenvolvimento satisfatório de suas respectivas trajetórias acadêmicas fortalecendo os seus processos formativos.

Através deste Edital, obtivemos a **aprovação de 02 (dois) monitores bolsistas**, Jaiane Moreira e Allane de Oliveira, que desenvolverão atividades remuneradas até setembro de 2023.

**(iii) Edital 24/2022** – Financiamento de projetos de extensão

A proposta foi **aprovada dentro do orçamento máximo para projetos de extensão, no valor de R\$7.400,00** (sete mil e quatrocentos reais), sendo R\$ 2.400,00 (dois e quatrocentos mil reais) direcionados ao pagamento de **01 (um) bolsista** (Cauê Oliveira) e o restante para materiais de consumo e pagamento de pessoas físicas e jurídicas. O Programa de Apoio a Projetos de Extensão (PROAPEX) é um instrumento que visa reafirmar a Extensão Universitária como processo acadêmico e institucionalizado através do apoio às ações desenvolvidas no âmbito da UNEB. Em consonância com o Plano de Desenvolvimento Institucional (PDI), são considerados os aspectos formativos e o compromisso social destas ações, visando o desenvolvimento regional e o fortalecimento da extensão desta.

**(iv) Edital 019/2022** – Bolsa de Iniciação Científica Júnior (IC-Júnior)

Obtivemos a **aprovação de 01 (um) monitor bolsista de Iniciação Científica Júnior** (Yaskara dos Santos). Este Programa busca promover a experiência científica para estudantes da Educação Básica da Rede Pública do Estado da Bahia, através da participação nas iniciativas de pesquisa e/ou tecnológicas, na Universidade do Estado da Bahia (UNEB). Visa-se, com este Programa, estimular a pesquisa, o desenvolvimento tecnológico e a inovação no âmbito da Educação Básica.

**(v) Edital 018/2022** – Bolsa de Iniciação Científica (IC)

Obtivemos a **aprovação de 01 (um) monitor bolsista de Iniciação Científica PICIN/UNEB** (Marina Felipe Barreto, subcoordenadora) e **01 (um) monitor voluntário** (Joanderson Borges, subcoordenador). O Programa de Iniciação Científica e o de Iniciação Tecnológica são instrumentos eficazes na formulação e aplicação da política de pesquisa da UNEB, bem como na qualificação de

estudantes de Cursos de Graduação, colaborando no desenvolvimento científico-tecnológico e regional.

**(vi) Edital 042/2021** – Financiamento de projetos de extensão

A proposta foi **aprovada dentro do orçamento máximo para projetos de extensão, no valor de R\$ 5.200,00** (cinco mil e duzentos reais), sendo R\$ 2.000,00 (dois mil reais) direcionados ao pagamento de **01 (um) bolsista** (Nalanda dos Anjos, que permanece no projeto na condição de voluntária) e o restante para materiais de consumo e pagamento de pessoas físicas e jurídicas.

**(vii) Edital 012/2021** – Projeto de Extensão

O Edital visa a concessão de Bolsas de Iniciação à Extensão a estudantes devidamente matriculados nos cursos de graduação da UNEB. Através dele, obtivemos a **aprovação de 02 (dois) monitores bolsistas**, Joanderson Borges e Haiane Pereira, que hoje permanecem nas subcoordenações, enquanto monitores voluntários.

**(viii) Publicações**

- SILVA, A. P. da; LEMOS, D. M.. Variável concordância verbal de terceira pessoa do plural: contribuições para o perfil sociolinguístico da comunidade de remanescentes quilombolas de Lagoinha. In: SILVA, J.; LEMOS, D.. **Educação, Linguística e Literatura: múltiplas faces do ensino**. Tutoiá/MA: Editora Diálogos, 2022.
- ALMEIDA, A. R. D de; LEMOS, D. M.. Variação linguística entre o “nós” e o “a gente” na comunidade quilombola lagoinha, São Gabriel-BA. In: SILVA, J.; LEMOS, D. M. (org.). **Estudos Linguísticos: (novos) caminhos**. v. 2. Tutoiá/MA: Editora Diálogos, 2021.
- LEMOS, D. M.; LACERDA, M. F.; PEREIRA, H. S.; BORGES, J.de A. **Notíciassobre o projeto ELIHS/UNEB: amostras de fala de comunidades rurais afro-brasileiras do território de irecê**. Salvador: EDUNEB, 2022 (no prelo).

O trabalho desenvolvido até o momento, no âmbito da pesquisa, tem ressaltado o caráter heteróclito da língua, especificamente ao se tratar do português popular brasileiro. Dessa forma, reafirmamos a importância de realização de pesquisas de descrição e análise do

português falado em comunidades afro-brasileiras no TII, variedades do português até então não catalogadas para um possível estudo sistemático no campo da linguística. Além do mais, não podemos deixar de relatar e compreender o compromisso social da Universidade (em especial a UNEB – *Campus XVI*) que está sediada neste território. A partir da composição e disponibilização do *corpus* do Projeto ELiHS, será possível entrelaçar, de forma mais efetiva, o papel do ensino, pesquisa e extensão, ao articular o conhecimento científico ao regional, permitindo a caracterização da identidade linguística de seus falantes.

É sabido que a compreensão da realidade social e linguística desse território do interior baiano, bem como de sua formação histórica, constitui um vasto campo de estudos ainda a ser, acertadamente, investigado por estudiosos vários. Se se pensar nas significativas diferenças entre as variedades linguísticas por cá faladas e reconhecidas pelos linguistas como um resultado de processos sucessivos de contatos e mudanças, veremos que, a partir daí, descortina-se um campo sócio-histórico, no qual afloram, para além do plano formal da estrutura linguística, perquirições sobre as circunstâncias sociais que irromperam tais variações no cenário do Território de Irecê.

Como já citado, Mattos e Silva (2008) faz referência à importância de extrair argumentos de dados de demografia histórica, de mobilidade sociocultural, geográfica, político-econômica, como fatores relevantes na recomposição de quadros sociolinguísticos; para tal feito, esperamos contar com amostras confiáveis. Torna-se, por conseguinte, indispensável ao leitor/pesquisador conhecer minuciosamente os critérios metodológicos adotados na organização do *corpus*, como propomos nas seções que seguem.

#### 4.2.2 Da estratificação à disponibilização das amostras: abordagens metodológicas

Bancos de dados linguísticos de fala, como este apresentado, têm favorecido a descrição do português brasileiro, numa proposta de interiorização; no entanto, nesse percurso, surgem desafios que envolvem, de forma pontual, a padronização de procedimentos metodológicos, no âmbito das pesquisas de orientação sociolinguística variacionista, visto que a comparação entre resultados obtidos de fenômenos linguísticos variáveis é um recurso analítico importante, e só a padronização metodológica pode favorecer uma comparação confiável. Nesse contexto, Freitag, Martins e Tavares (2012, p. 918, grifo nosso) afirmam que

a cada projeto que constitui seu banco de dados em uma comunidade de fala, o mapeamento das variedades do português no Brasil vai se efetivando, *mas só a*



*padronização dos procedimentos metodológicos permitirá a realização de estudos contrastivos entre as variedades, para, então, possibilitar uma descrição mais acurada do português brasileiro.*

Os autores seguem afirmando que é possível observar que muitas pesquisas no campo da sociolinguística variacionista apontam importantes resultados, mas, no que tange à abordagem metodológica, restringem-se a citar como o *corpus* foi selecionado, com quantos participantes<sup>49</sup>, qual o banco de dados, e, por fim, a estratificação das células sociais. O breve relato, direcionado às descrições metodológicas em trabalhos científicos, não condiz com a dinâmica e esforços dispensados na constituição, organização e disponibilização de banco de dados, que perpassa por várias etapas (*cf.* Figura 26).

**Figura 26** – Metodologia para organização de *corpus*



Fonte: Elaborada pela autora.

As sucessivas etapas postas em uma pesquisa que objetiva a constituição de amostras de fala envolvem cuidados que vão desde o estabelecimento de critérios metodológicos à necessidade de recursos humanos e financeiros. A última etapa, e não menos importante, diz respeito à disponibilização do acervo, em que o Projeto cumpre seu papel ético e social, visto que todo o trabalho realizado é fruto de um financiamento através de recursos públicos, como descrito na seção anterior. Nesse campo de armazenamento e disponibilização de dados, podemos citar: Para História do Português Brasileiro (PHPB)<sup>50</sup>; Corpus do Laboratório de História do Português<sup>51</sup>; Programa para a História da Língua Portuguesa (PROHPOR)<sup>52</sup>; Corpus Eletrônico de Documentos Históricos do Sertão (CE-DOHS)<sup>53</sup>; Corpus Histórico do português Tycho Brahe<sup>54</sup>, dentre outros. Se observamos esses Projetos citados, claro que não é uma lista exaustiva, muitos disponibilizam, em sua maioria *corpus* escritos, um indicativo da

<sup>49</sup> Atualmente o Comitê de Ética em Pesquisa indica a utilização do termo *participantes*, ao pensarmos nas pesquisas que envolvem seres humanos, visto que esses são atuantes e participativos no processo de constituição de amostras, não apenas um informante passivo, a ser detalhadamente descrito aos olhos da ciência. Por isso, assumimos o termo *participante*, compreendendo seu papel ativo na constituição das amostras do ELiHS.

<sup>50</sup> Conferir site: <https://sites.google.com/site/corporaphpb/>

<sup>51</sup> Conferir site: <https://laborhistorico.letras.ufirj.br/>

<sup>52</sup> Conferir site: <https://www.prohpor.org/bit-prohpor>

<sup>53</sup> Conferir site: <http://www5.uefs.br/cedohs/view/home.html>

<sup>54</sup> Conferir site: <http://www.tycho.iel.unicamp.br/corpus/>

necessidade de investirmos mais na organização, armazenamento e disponibilização de *corpora* orais, não se fechando aos Projetos e suas respectivas Instituições.

Que possamos, a partir da disponibilização de novos acervos, fazer “a arte de fazer bom uso de dados ruins” (Labov, 1994, p. 11), visto que nem sempre dispomos de dados de qualidade e nem sempre os dados são facilmente acessíveis (Freitag *et al*, 2021). Sendo assim, comungando com a ideia de que a abordagem metodológica perpassa pelo prisma da pesquisa Sociolinguística, torna-se importante compreender que a idealização de uma unificação/homogeneidade metodológica trará confiabilidade aos dados, evitando possíveis divergências nos resultados das pesquisas e que “o trabalho com *corpus* não é ‘*commodo*’ – pelo contrário, é por vezes penoso, sempre curioso, às vezes divertido [...]. Sem dúvida é um trabalho necessário, como base para a reconstrução do passado linguístico do português que aqui se formou, o português brasileiro” (Mattos e Silva, 2004a, p. 120). Assim, propomo-nos, aqui, a apresentar a metodologia e os pressupostos norteadores da constituição da amostra sincrônica do banco de dados do ELiHS.

#### 4.2.2.1 Comitê de Ética em Pesquisa

Iniciaremos pela importância do Comitê de Ética, visto que, na passagem do século XX para o século XXI, as pesquisas que lidam diretamente com seres humanos<sup>55</sup>, mesmo que no âmbito das linguagens, passaram a ser condicionadas à aprovação pelo Comitê de Ética em Pesquisa (doravante CEP), que tem imposto algumas restrições no desenvolvimento e execução de projetos, afinal “[...] há riscos na documentação sociolinguística, e que podem e precisam ser minimizados, com a elaboração prévia do desenho da pesquisa, com o treinamento adequado do pesquisador de campo e com o cuidado na divulgação dos resultados [...]” (Freitag, 2017, p. 10).

Dentre as situações de riscos a que podem estar expostos os participantes que colaboram com uma coleta de dados de cunho sociolinguístico, Freitag (2017) destaca: exposição da imagem e voz; exposição de informações pessoais; constrangimentos causados pela abordagem na entrevista; reativação de memórias ou situações vividas que causam sofrimento psíquico.

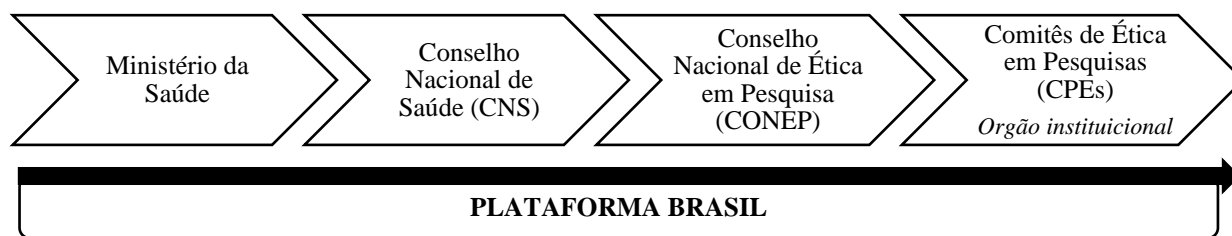
---

<sup>55</sup>A Resolução nº 466/2012, que aprova as diretrizes e normas regulamentadoras de pesquisas envolvendo seres humanos, define que a pesquisa envolvendo seres humanos refere-se a “[...] pesquisa que, individual ou coletivamente, tenha como participante o ser humano, em sua totalidade ou partes dele, e o envolva de forma direta ou indireta, incluindo o manejo de seus dados, informações ou materiais biológicos”.

Sendo assim, cabe a ressalva para o fato de este Projeto assumir o compromisso diante das prerrogativas da Resolução CNS 510/2016, tendo anuência, principalmente, do seu Art. 3 parágrafo VII, que garante a confidencialidade das informações, da privacidade dos participantes e da proteção de sua identidade, inclusive do uso de sua imagem e voz, durante todas as fases da pesquisa.

Para conhecimento, a aprovação de Projeto que envolve seres humanos é avaliada pelo CEP, que se responsabiliza pelo acompanhamento e aprovação, obedecendo a estrutura hierárquica descrita na *Figura 27*.

**Figura 27** – Estrutura hierárquica de responsabilização por pesquisas envolvendo humanos no Brasil



Fonte: Elaborada pela autora.

No campo da Sociolinguística, as discussões e os resultados de descrições empíricas de padrões linguísticos surgem dos levantamentos realizados em *corpora*, o que nos faz persistir e caminhar na tentativa de atender a todos os requisitos, compreendendo a importância da regularização dos projetos que envolvem seres humanos, bem como visibilidade científica dada através da sua aprovação. Para registro, a proposta desta tese possui o Certificado de Apresentação para Apreciação Ética (CAAE) número 39746320.1.0000.0053.

#### 4.2.2.2 Seleção das comunidades e dos participantes

Na maioria dos trabalhos em Dialectologia e Sociolinguística Variacionista, a escolha do lugar em que as amostras serão coletadas enquadra-se nas questões metodológicas, visto que esses lugares são representados por isoglossas, as quais permitem traçar características dialetais do Brasil, prefigurando a identidade linguística dessas comunidades.

A partir dessa perspectiva, os falantes são vistos como construtores de lugar, na medida em que experimentam o espaço físico e social. Esse lugar, para Sociolinguística, configura-se para além de uma simples localização geográfica, haja vista que perpassa pelas questões sociais

e históricas que marcam esses lugares e entrelugares, um contexto significativo da ação humana, em que a língua se apresenta heteróclita e multifacetada.

Como já descrito, este Projeto tem suas amostras concentradas no Território de Identidade de Irecê, e, quando falamos em Território de Identidade, não estamos falando apenas de conceito político-administrativo, limitado às fronteiras e tendo o estado como única fonte de poder no controle espacial. Falamos no conceito de território em outras abordagens, a de territorialidade, a qual designa a qualidade que o território ganha de acordo com a sua utilização ou apreensão pelo ser humano, a partir da análise de processos econômicos, políticos e culturais que se territorializam no tempo e no espaço, o que suscita compreender o território com outras possibilidades de constituição de outras territorialidades, com base em dimensões até então não experimentadas (Saquet, 2009). Sim, estamos falando de um espaço socialmente construído, das relações históricas, sociais e culturais, de uma dinâmica própria, das nossas comunidades de fala.

Cabe essa retomada de discussão, iniciada na *Seção 4.1.4*, visto que, dentre os procedimentos metodológicos iniciais, destaca-se a seleção das comunidades e dos participantes da amostra, fundamentais para a constituição de banco de dados, especialmente de fala espontânea, como a ser disponibilizada através deste Projeto. Como já discutido na seção anterior, tendo um *corpus* como este aqui apresentado, e cientes das discussões em torno da Ciência Aberta<sup>56</sup> e da idealização de critérios metodológicos padronizados, não podemos nos isentar de tais demandas metodológicas; por isso, as decisões aqui tomadas, especialmente para a coleta das amostras de fala, demandaram planejamento, com definição de critérios de estratificação de amostras que fossem claros e objetivos, possíveis de serem replicáveis em outros contextos sociolinguísticos. Assim, seguiremos apresentando esses critérios que marcam a representatividade dessas comunidades e desses falantes.

Como já apresentado na *Seção 4.1.3*, o Território de Irecê é culturalmente e linguisticamente contemplado por aproximadamente 148 (cento e quarenta e oito) comunidades afro-brasileiras, certificadas pela Fundação Palmares enquanto Remanescentes de Quilombos, contexto desafiador para uma, inicialmente, pequena estratificação de amostras. Portanto, critérios foram estabelecidos, na tentativa de parametrizar as escolhas, não perdendo o foco na representatividade das amostras.

---

<sup>56</sup> Como afirma Freitag (2021; 2017), a documentação sociolinguística é realizada no Brasil para alimentar descrição linguística, por isso, é importante garantir a adoção de padrões que permitam a comparabilidade de resultados, sem possíveis tendenciosismos. Tais padrões metodológicos devem atender as diversas demandas a ponto de alcançarmos a possibilidade de sua replicabilidade e atender aos critérios exigidos pela perspectiva da Ciência Aberta, a qual busca disponibilizar dados com licenças de uso claros e acessíveis, trabalho desenvolvido pelos nossos Comitês de Ética em Pesquisa (CEPs).

Para o dimensionamento do *corpus*, foram selecionadas 06 (seis) cidades representativas do Território, que possuíam grau de urbanização menor que a média estadual de 72,1% (*cf.* Gráfico 3) e que estavam em torno de Irecê, cidade sede da UNEB – *Campus XVI*. Esse dado é importante, pois os processos de urbanização e de escolarização podem minimizar as marcas linguísticas geradas pelo contato entre povos, assim como se espera para essas comunidades a comporem a amostra do Projeto ELiHS; por isso, selecionamos, a partir dos dados estatísticos, as cidades que possuem menor grau de urbanização.

Na sequência, apresentaremos mais três critérios de refinamento das amostras, relativos às comunidades, a saber:

- (i) Com Certificação na Fundação Palmares;
- (ii) Processo de regularização aberto no INCRA;
- (iii) Dados adquiridos diante de formulário eletrônico (*Google Forms*), enviado a membros representativos das Comunidades do Território.

O critério (i) fundamenta-se na Fundação Cultural Palmares, criada no dia 22 de agosto de 1988, pelo Governo Federal, a qual se configura como “a primeira instituição pública voltada para promoção e preservação dos valores culturais, históricos, sociais e econômicos decorrentes da influência negra na formação da sociedade brasileira” (*site FCP*, 2021). Em 16 de fevereiro de 2023, a FCP havia emitido 148 certificações no Território de Irecê (*cf.* Quadro 8), conferindo acesso aos programas sociais do Governo Federal. No entanto, a partir do Decreto nº 4.887, de 20 novembro de 2003, a competência de regularização fundiária passou a ser do Instituto Nacional de Colonização e Reforma Agrária (INCRA).

Após uma exaustiva busca de informações mais específicas, que marcasse a representatividade das comunidades certificadas, fomos surpreendidos por uma segmentação de informações que são fechadas às associações de moradores das comunidades. As lacunas deixadas para o prosseguimento de uma pesquisa de cunho sociolinguístico nos fizeram compreender a necessidade de iniciamos essa caracterização das Comunidades do Território, através do formulário eletrônico (iii), planejado e executado na tentativa de reunirmos informações que colaborassem na caracterização dessas Comunidades, favorecendo o traçado de estratégias metodológicas mais ajustadas à estratificação das amostras que devem compor o banco de dados do Projeto ELiHS.

Dando prosseguimento à estratificação das amostras, estabelecemos novos critérios de refinamento, diante dos resultados gerados pelo questionário descrito no *Quadro 7*, especificamente a tematização das questões:

- Q. 9 – Tempo de formação;
- Q. 10 – Número de membros;
- Q. 11 – Etnicidade;
- Q. 14 e 15 – Informatização/acesso à internet e ao telefone;
- Q. 16 e 178 – Nível de escolarização;
- Q.18 – Distanciamento da sede (centro urbano).

Assim, após a análise dos resultados, chegamos aos seguintes critérios sociodemográficos para refinamento da amostra, considerando os critérios (i) e (ii), já apontados, que reduziram nossa amostra a 70 (setenta) Comunidades, visto que apenas essas atentem aos critérios (i) e (ii). Essas foram escalonadas considerando as comunidades caracterizadas por serem:

- [ - ] Urbanas (estatística do IBGE e distanciamento da sede);
- [ + ] Afrodescendente (autodeclaração – a partir de 90% dos membros);
- [ - ] Informatizada (autodeclaração – sem ou difícil acesso à internet);
- [ - ] Escolarizada (toda a amostra será representativa do português popular).

Os critérios supramencionados possibilitam a comparabilidade de resultados (Freitag, 2018, 2021), já que a estratificação da amostra foi parametrizada em categorias sociodemográficas amplas, que garantisse a representatividade esperada, em um universo de mais de 148 comunidades afro-brasileiras existentes no Território de Irecê. A possibilidade de escolhermos apenas 01 (uma) comunidade para compor a amostra do Projeto daria margem ao questionamento dos dados/resultados serem representativos do Território ou apenas daquela comunidade estudada, por isso a seleção inicial de 06 (seis) Comunidades<sup>57</sup> que integram, inicialmente, o *corpus* do ELiHS traz uma proposta quantitativa e, principalmente, qualitativa, com contextos históricos e socioeconômicos diversos – alinhados aos critérios de estratificação.

---

<sup>57</sup> Das 06 (seis) Comunidades, apenas Volta do Angico, uma das mais antigas comunidades, comporá a amostra para análise do fenômeno variável da concordância nominal de número.

**Quadro 11** – Características sociodemográficas das Comunidades que comporão o *corpus* do Projeto ELiHS

COMUNIDADE/CIDADE	Tempo de formação*	Nº de membros*	Etnicidade negra	Informatização/ acesso à internet e ao telefone	Distanciament o da sede
Volta do Angico – Canarana	± 200 anos	± 350	90%	Difícil acesso	19km
Lagoinha – São Gabriel	± 95 anos	± 350	100%	Difícil acesso	38km
Lapinha – América Dourada	± 120 anos	±140	90%	Difícil acesso	17km
Sertão Bonito – João Dourado	±73 anos	±70	100%	Não possui acesso à internet e difícil acesso ao telefone	70 km
Lagedinho – Lapão	± 80 anos	±196	100%	Difícil acesso	36 km
Serra Grande– Ibipeba	± 100 anos	± 140	100%	Difícil acesso	24 km

\* As Comunidades não sabem informar o tempo exatado de formação e quantitativo de membros, por isso a utilização do símbolo (±).

Fonte: Elaborado pela autora.

Os critérios expostos no *Quadro 11* foram estabelecidos diante da inadequada apropriação do termo de [+ isolada], outrora utilizado por Lucchesi e Baxter (2009) para as comunidades rurais afro-brasileiras retratadas em suas pesquisas. Diante da expansão das tecnologias de comunicação e informação, torna-se metodologicamente inviável utilizarmos tal critério no processo de estratificação, visto que grande parte das comunidades têm acesso, em maior ou em menor grau, aos meios de comunicação de massa (*cf.* Gráfico 10 e 11) e aos centros urbanos. Por isso, nesse critério, optamos por tratar nos termos do continuum proposto por Bortoni-Ricardo (2021), que está entre as redes isoladas e redes integradas (*cf.* seção 4.12)

Ainda vale salientar que o *corpus* constituído preconiza um processo de amostragem não aleatória, visto que, para a Sociolinguística Variacionista,

[...] a quantidade de falantes das categorias (células sociais) é pré-definida, e o pesquisador de campo vai em busca de falantes disponíveis a participar como voluntários do processo de entrevista sociolinguística (que demanda certo tempo), que sejam representativos da comunidade de fala (que não causem estranhamento, que não destoem do padrão da comunidade, etc.). *A amostragem não é, portanto, aleatória, pois não são todos os falantes da população que têm igual chance de ser selecionados para a amostra* (Freitag, 2018, p. 671-672, grifo nosso).

Como aponta Freitag (2018), é preciso cautela na organização e estratificação da amostragem não aleatória, já que vieses amostrais (e também de publicação<sup>58</sup>) podem levar a

<sup>58</sup> “[...] o viés de publicação também merece reflexões: achados positivos têm maior probabilidade de ser publicados do que achados negativos. Na abordagem da sociolinguística variacionista, esse viés se manifesta

resultados equivocados, incidindo diretamente na representatividade da amostra. Numa releitura de Tagliamonte (2006), Freitag (2018) aponta:

- (i) Influência humana da escolha (sentimentos, afinidades e atitudes);
- (ii) Inabilidade para encontrar certos segmentos da população;
- (iii) Cobertura inadequada da população amostra;
- (iv) Falta de cooperação de alguns subgrupos.

Para (i), fomos cuidadosos ao inserirmos, na Ficha do Participante, dados que indicam o grau de relação<sup>59</sup> entre documentador e participante (*cf.* Quadro 12).

**Quadro 12** – Ficha do participante: influência humana da escolha

<b><u>DA ENTREVISTA</u></b>	
Data da Entrevista:	_____ Horário: _____
Quantidade de horas/minutos gravados:	_____
Local da Entrevista:	_____
Presença e atuação de terceiros:	( ) Não. ( ) Sim. Quais? _____
Qual o grau de relação entre informante e documentador:	
( ) Grau 1 – Bastante próximo. Os falantes têm laços fortes (amizade, parentesco, colega de trabalho ou escola etc.) e interagem diariamente;	
( ) Grau 2 – Próximo. Os falantes interagem frequentemente, mas não têm laços fortes;	
( ) Grau 3 – Próximo. Os falantes não interagem frequentemente e não têm laços fortes;	
( ) Grau 4 – Neutro. Os falantes se conhecem, mas não interagem com frequência;	
( ) Grau 5 – Distante. Os interlocutores não se conheciam anteriormente e só conversaram no momento da gravação da interação.	

Fonte: Elabora pela autora.

Esse cuidado se deu, principalmente, pelo difícil acesso às comunidades, que ainda se mantêm fechadas a membros externos, sendo as nossas entrevistas, em sua maioria, direcionadas/realizadas por alunos/monitores quilombolas, para que o enviesamento proposto em (iv) seja, ao menos, neutralizado.

Em relação ao (ii), relatamos dificuldades em preenchermos as células direcionadas aos informantes da Faixa I (até 25 anos), já que a amostra é representativa do português popular e essas devem ser preenchidas com informantes que cursaram até o 5º ano do Ensino

---

quando os pesquisadores omitem/não informam as variáveis extralinguísticas (preditoras) que foram controladas no modelo, mas que não apresentaram significância estatística. O fato de um fator não ser estatisticamente significativo é sociolinguisticamente significativo!” (Freitag, 2018, p. 672).

<sup>59</sup> Sugestão de Freitag (2018, p. 679), “com base na proposta de Blake e Josey (2003), Oushiro (2011) e Araújo, Santos e Freitag (2014) desdobram-se critérios para controle de densidade e complexidade da rede de falantes”.



Fundamental. Esse é um desafio apontado por outros autores: “[...] com a quase universalização da educação básica e o crescimento exponencial na educação superior, é tarefa difícil preencher determinadas células sociais de bancos de dados sociolinguísticos, *como a de um jovem analfabeto, mesmo fora dos grandes centros*” (Freitag, 2017, p. 16).

Todos esses critérios e cuidados, já mencionados, apontam uma preocupação com a representatividade da amostra, visto que “há um viés de intencionalidade para tornar o processo de seleção de falantes compatível com o construto da população: a comunidade de fala” (Freitag, 2018, p. 673, grifo nosso).

Em relação a (iii), a seleção dos participantes da amostra se dará por cotas proporcionais (Freitag, 2018, p. 675), que computará 02 (dois) informantes, para cada célula social, para Comunidades com mais de 200 (duzentos) membros, e 01 (um) informante, para cada célula, para Comunidades com até 200 (duzentos) membros. A distribuição por cotas proporcionais implica amostras que mais se aproximam da realidade populacional das comunidades, em termos de proporcionalidade (*cf.* Quadro 13).

**Quadro 13** – Quadro amostral

COMUNIDADES	NÚMERO DE HABITANTES	NÚMERO DE PARTICIPANTES	ENTREVISTAS GRAVADAS	ENTREVISTAS TRANSCRITAS
Volta do Angico	± 350	12	12	12
Lagoinha	± 450	12	06	06
Lapinha	±140	06	00	00
Sertão Bonito	±70	06	00	00
Lagedinho	±196	06	00	00
Serra Grande	± 140	06	00	00

Fonte: Elabora pela autora.

Em relação a (iv), as parcerias traçadas com as centrais quilombolas e conselho territorial (*cf.* seção 4.2.3) colaboraram na cooperação das comunidades para desenvolvimento da pesquisa, envolvendo: reconhecimento do espaço, levantamento de dados, seleção de participantes, gravações e registros fotográficos dos espaços públicos.

Sendo assim, reafirmamos a importância de critérios metodológicos, na tentativa de que a amostra, composta no âmbito do Projeto ELiHS, siga as orientações da Sociolinguística Variacionista, buscando uma padronização que atenda à expectativa de ser representativa da

diversidade das comunidades afro-brasileiras do Território de Irecê. Assim, a estratificação da nossa amostra considera fatores sociodemográficos que atuam sobre a língua, a saber:

- (i) Grau de Escolarização – Não escolarizados ou que estudaram até o 5º ano do Ensino Fundamental<sup>60</sup>;
- (ii) Nascidos e residentes na Comunidade;
- (iii) Sexo de registro – Masculino e Feminino;
- (iv) Faixa Etária<sup>61</sup> – I (20-39), II (40-59), III (acima de 60 anos)<sup>62</sup>.

Quando finalizado, o banco de dados contará com 48 (quarenta e oito) gravações de áudio, com suas respectivas transcrições grafemáticas, em formato PDF, e em edição xml, acessíveis gratuitamente através do *site* do Projeto. Em relação ao processo de gravação, a seção 4.2.2.3 tratará de suas especificidades.

#### 4.2.2.3 Gravação

Como já descrito, a metodologia de estratificação da amostra prevê que os potenciais participantes passem por dois momentos, a saber:

- (i) A pré-seleção, com preenchimento da Ficha social do participante para checagem do seu perfil sociocultural e enquadramento aos critérios da amostra;
- (ii) Realização da entrevista sociolinguística, a partir de um roteiro pré-estabelecido;

O (i) antecede a entrevista propriamente dita e tem por objetivo norteador a obtenção de informações relevantes sobre o participante, a saber: idade, escolaridade, ocupação (atividade profissional), contexto familiar. Já em (ii), utilizaremos o método DID (Diálogo entre informante e Documentador), com duração média de 30 a 60 minutos, seguindo o modelo de entrevista sociolinguística semidirigida, em que entrevistado e entrevistador conversarão de

---

<sup>60</sup> A inserção da prerrogativa do informante ter estudado até o 5º ano de Ensino Fundamental se dá pelo fato da dificuldade de preenchimento da célula da Faixa I, visto o avanço no processo de escolarização em comunidades afastadas – Educação do campo, EJA e outros.

<sup>61</sup> Como afirma Lucchesi (2009, p. 155), as amostras com falantes com mais de 80 anos permitem visualizar os estágios mais antigos da gramática das comunidades de fala analisadas com o intuito de se identificar os traços mais ligados aos efeitos do contato entre línguas que marca a origem dessas comunidades.

<sup>62</sup> A estratificação etária dos informantes segue a padronização do Instituto Brasileiro de Geografia e Estatística – IBGE, computando 04 (quatro) faixas (20-39; 40-59; acima de 60 anos. Iniciamos as gravações que comporão a Faixa IV, acima de 80 anos, mas essa não foi finalizada pela dificuldade em encontramos participantes nesta faixa etária vivos e lúcidos.

forma livre e espontânea sobre temas diversos, a critério do participante entrevistado. No entanto o entrevistador terá em mãos um roteiro de entrevista que o auxiliará na condução da conversa/entrevista. Os temas circunscrevem-se nos seguintes eixos: Pessoal; Infância; Família; Trabalho/ocupação; Lazer; Cidade/comunidade; Autorreconhecimento enquanto comunidades afro-brasileiras.

Com afirma Freitag (2017, p. 23), “para possibilitar uma amostra de fala que seja representativa da performance do indivíduo, a situação de entrevista precisa desencadear a produção de diferentes tipos textuais/sequências discursivas”. O roteiro pré-elaborado, portanto, não se configura como um questionário, visto que contempla assuntos que permitem narrativas pessoais e possibilita ao entrevistador explorar diferentes tipos de sequências discursivas. Sendo assim, como aponta Tarallo (2007, p. 23), “a narrativa de experiência pessoal é a mina de ouro que o pesquisador-sociolinguista procura”.

Todos esses critérios já apontados, e outros a serem descritos, favorecem a coleta de uma fala menos monitorada; no entanto não podemos deixar de relatar que, ao realizarmos entrevistas sociolinguísticas, esbarramo-nos no que Labov (2008 [1972]) nomeia como “paradoxo do observador” – há um desejo de observar a fala do indivíduo quando ele não é observado; utilizamos, todavia, gravadores e seguimos os protocolos sugeridos pelo Comitê de Ética em Pesquisa.

No contexto do “paradoxo do observador”, sabe-se que a presença do gravador interfere na naturalidade da situação comunicativa, porém “o pesquisador deverá tentar neutralizar a força exercida pela presença do gravador e por sua própria presença como elemento estranho à comunidade” (Tarallo, 2007, p. 21). Temos, dessa forma, investido em algumas alternativas que têm favorecido a neutralidade de paradoxo apontado por Labov, são elas:

- Treinamento dos entrevistadores<sup>63</sup> – com simulações em pares;
- Realização das gravações por membros quilombolas – representatividade étnica;
- Relatar que, além do objetivo linguístico, o projeto prevê o empenho na aquisição de políticas públicas para as comunidades<sup>64</sup>, bem como propostas de intervenções emergenciais;

---

<sup>63</sup> Orientados a partir dos pressupostos de Freitag (2017), em seu capítulo “Treinamento para a entrevista sociolinguística”, em que tece alguns esclarecimentos e técnicas que favorecem nosso maior objetivo – coletar amostras em contexto informal.

<sup>64</sup> A coordenadora Dayane Lemos tem participado das reuniões do Colegiado de Desenvolvimento Territorial (CODETER), apresentando os resultados preliminares e buscando alternativas emergenciais junto aos membros representativos do Território.

- Gravador de voz portátil de cor preta – discreto e de fácil manuseio;
- Realizar a gravação no dia, local e hora da conveniência do participante.

Reiteramos que o método utilizado no processo de gravação implica um diálogo semidirigido, em que o entrevistador tenta neutralizar o ambiente de gravação, no intuito de que o falante utilize o português em ambientes informais; portanto, “essa entrevista, longe de ser um questionário, deve-se constituir de uma conversa, a mais informal possível, apesar das circunstâncias, adversas, posto que o que se quer é a *fala casual, habitual, dos falantes*” (Silva, 2010, p. 125, grifo nosso). Não perdendo de vista que esse processo não se constitui como uma ação passiva, um observar ocasional, é preciso cautela na condução da observação e gravação.

#### 4.2.2.4 Transcrição

Os minutos de áudios foram submetidos à transcrição grafemática, na tentativa de, sempre que possível, buscar uma aproximação das características linguísticas dos falantes, demarcando marcas específicas do seu dialeto, possibilitando o mapeamento de fenômenos linguísticos, em especial os morfossintáticos.

A chave de transcrição busca a padronização do *corpus* oral do *Projeto Estudos Linguístico e Histórico do Sertão (ELiHS)*, apoiando-se na proposta de transcrição ortográfica<sup>65</sup> que orienta a transcrição dos inquéritos realizados no âmbito do *Projeto Vertentes do Português Popular do Estado da Bahia*, o qual também objetiva contribuir para o conhecimento da realidade linguística brasileira e, de forma específica, para o estudo da língua falada em áreas do semiárido baiano. No arquivo/chave de transcrição constam:

- Cabeçalho (*cf.* Quadro 14);
- Indicações gerais sobre a formatação do texto transcrito (*cf.* Quadro 15);
- Critérios gerais para orientar a transcrição (*cf.* Quadro 16).

---

<sup>65</sup> Disponível em: [http://www.vertentes.ufba.br/images/paginas/projeto/chave\\_de\\_transcricao.pdf](http://www.vertentes.ufba.br/images/paginas/projeto/chave_de_transcricao.pdf). Acesso em: 31 ago. 2020.

**Quadro 14** – Cabeçalho padrão da chave de transcrição

<b>DADOS DO PARTICIPANTE</b>
<b>NOME:</b> <i>garantir o sigilo desta informação – utilização de siglas nominais</i>
<b>SEXO DE REGISTRO:</b> ( ) M ( ) F <b>IDADE:</b>
<b>PROFISSÃO:</b>
<b>COMUNIDADE:</b>
<b>TEMPO DE GRAVAÇÃO:</b>

Fonte: Elaborada pela autora.

**Quadro 15** – Indicações gerais sobre a formatação do texto transcrito

<b>CRITÉRIOS GERAIS</b>	
Trecho ininteligível	[ININT]
Interrupção do inquérito	[INTERRUP]
Trecho sobre qual não há certeza na audição	PART: A comunidade [surgiu] faz tempo
Comentário durante a fala do informante, sem interromper o fluxo sintático-discursivo, deve-se indicar com reticências a continuidade desse fluxo do informante, intercalando a fala do documentador	PART: A comunidade [surgiu] faz tempo... DOC: Ah, legal! PART: ... mas só agora temos o registro.
Discurso direto, deve-se utilizar aspas e pontuar antes de fechar as aspas	PART: Ele disse: “Eu vô embora.”

Fonte: Elaborada pela autora.

**Quadro 16** – Critérios gerais para orientar a transcrição

<b>CRITÉRIOS PARA TRANSCRIÇÃO ORTOGRÁFICA</b>		
<b>Nível morfosintático</b>	Variante zero na concordância nominal variável	As comunidade; Estava no comunidade.
	Concordância verbal variável	As comunidade é unidas. Nós vai para o centro.
	Omissão de preposições, artigos, complementizadores, etc.	Deu Maria isso aí.
	Repetições, correções, hesitação, utiliza-se as reticências	As pessoas que...que vêm aqui. E as... Todos se divertem por aqui É... é... é... sono felizes.
	Separar o tópico por vírgula	Aquelas crianças, nós cuidaremos delas.
<b>Nível fônico</b>	<b>NÃO DEVEM SER REGISTRADOS</b>	
	Elevação das vogais médias em distribuição pré-acentuada, quer no interior de um mesmo vocábulo, quer no interior de um grupo de força (como no caso das sequências com a preposição <b>em</b> ), e pós-acentuada não-final.	<i>estava, e não istava; domingo, e não dumingo; em casa, e não im casa; rédea, e não rédia.</i>

Elevação das vogais médias em posição átona final.	nomi> <i>nome</i> ; matu> <i>mato</i>
Ditongação antes de consoante constrictiva implosiva.	[meys] > <i>mês</i>
Ditongação da terminação nasal -em	<i>tem</i> , e não <i>teim</i> ; <i>ninguém</i> , e não <i>ninguéim</i> .
Palatalização de <b>t</b> e <b>d</b> , antes de vogal palatal <sup>66</sup>	<i>ponte</i> , e não <i>pontche</i> ; <i>noite</i> , e não <i>noitche</i> .
Epêntese da vogal alta que desfaz o travamento de grupo consonântico <sup>67</sup>	<i>advogado</i> , e não <i>adivogado</i> ; <i>psicologia</i> , e não <i>ñopissicologia</i>
Vocalização da consoante lateral pós-vocálica	<i>mel</i> , e não <i>méu</i> ; <i>Brasil</i> , e não <i>Brasiu</i> ; <i>animal</i> , e não <i>animau</i>

Fonte: Elaborada pela autora.

Esse material, baseando-se no Projeto Vertentes, busca facilitar trabalho de transcrição, identificando visivelmente os fatos que devem, ou não, ser registrados na transcrição, lembrando que o objetivo dessa transcrição “[...] é transpor o discurso falado, da forma mais fiel possível, para registros gráficos mais permanentes, necessidade que decorre do fato de que não conseguimos estudar o oral através do próprio oral” (Paiva, 2010, p. 136), tarefa que não pode ser feita arbitrariamente, mas com fundamentos e posicionamentos teóricos que norteiem a transcrição.

A partir da transcrição grafemática, optamos por utilizar, concomitante, a edição XML (*Extensible Markup Language*), uma vez que, através desse formato,

é possível buscar aproximações entre as áreas, com a padronização ortográfica de corpora sociolinguísticos e o *desenvolvimento de etiquetas XML* para codificar a informação social de modo *a ser processável por softwares concordanceadores*. Para possibilitar a manipulação dos dados, a transcrição ortográfica padrão é mais eficiente [...] (Freitag, 2013, p. 163, grifo nosso).

Tomando como referência os critérios metodológicos adotados pelo CE-DOHS, projeto parceiro, as transcrições seguem os princípios ortográficos da escrita do português<sup>68</sup>, o que nos permite fazer uso das ferramentas computacionais da Linguística de *Corpus*.

Ainda cabe relatar que a linguagem *xml* será processada via ferramenta de edição eletrônica *eDictor*<sup>69</sup>, que permite a manipulação eletrônica do *corpus*, por meio de anotação e do uso de ferramentas de buscas de dados linguísticos. Assim, teremos a opção de converter a

<sup>66</sup> Palatalizações mais desviantes da norma urbana devem ser registradas: *muntcho*, *dôdio*, *tchem*, *tcherra*.

<sup>67</sup> A inserção de uma vogal média deve ser registrada: *adevogado*.

<sup>68</sup> A padronização favorece a utilização de ferramentas computacionais que buscam a otimização dos dados.

<sup>69</sup> Ferramenta desenvolvida por Maria Clara Paixão de Sousa, Fábio Kepler, Pablo Faria e, na versão *online*, por Luís Veronese.

versão em *word* para a versão em *xml*, não perdendo o arquivo que traz as transcrições grafemáticas, fidedignas das amostras de fala.

#### 4.2.3 A rede de pesquisa: parcerias importantes

Neste ponto, cabe frisarmos que o Projeto ELiHS integra um projeto maior, intitulado *Corpus Eletrônico de Documentos Históricos do Sertão (CE-DOHS)*, sediado no Departamento de Letras e Artes, da Universidade Estadual de Feira de Santana, e coordenado pelas profas. Dras. Zenaide de Oliveira Novais Carneiro e Mariana Fagundes de Oliveira Lacerda.

A Plataforma eletrônica do CE-DOHS pode ser acessada através do link de acesso <<http://www5.uefs.br/cedohs/>>, conforme *Figura 28*.

**Figura 28** – Plataforma eletrônica CE-DOHS

Fonte: Site CE-DOHS.

A *Figura 28* apresenta a face eletrônica gratuita do Projeto CE-DOHS, em que os *menus* possibilitam: Apresentação da proposta do Projeto; Equipe; Convênios; Parcerias; Coletâneas; Extensão; e Outros. O CE-DOHS apresenta uma base documental, majoritariamente inédita, para estudo do Português Brasileiro (PB), constitui-se em um *corpus* resultante de uma intensa prospecção em fontes metodológicas confiáveis. Suas coleções documentais são organizadas em: Edições em diferentes formatos; Edições com anotação sintática; Locais de produção e/ou recolha; Locais de nascimento; Sites. As Edições em diferentes formatos nos encaminham a um campo eletrônico de *corpora* que vai desde manuscritos datados de 1500-1982, a manuscritos, Impressos e Amostra de fala datados de 1823-2000.

**Figura 29** – Plataforma eletrônica CE-DOHS: *corpus* eletrônico

Feira de Santana - BA, Brasil  
Corpus Eletrônico de Documentos Históricos do Sertão [ CE-DOHS ]  
Voltar

1823-2000 | 1500-1822  
Manuscritos Impressos Amostras de fala Manuscritos

**Corpus Compartilhado Diacrônico – Manuscritos (PHPB-BA/Tycho Brahe/PROHPOR)**

Os documentos estão editados em XML, utilizando a ferramenta eDictor (desenvolvida por Pablo Faria, Fábio Kepler e Maria Clara Paixão de Sousa). Essa tecnologia de edição digital foi inspirada no *Corpus Histórico do Português Tycho Brahe*, coordenado por Charlotte Galves.

Coleções documentais/Edições em diferentes formatos	Número de documentos
Cartas para vários destinatários	208 Cartas

Coleção documental composta da edição fac-similar e semidiplomática de 208 cartas, extraídas de Carneiro (2005). Essa coleção, dividida em diversos destinatários, está depositada no Instituto Genealogia e

Projeto CE-DOHS  
Corpus Eletrônico de Documentos Históricos do Sertão  
FAPESB, Processo 5566/2010/Consepe: 202/2010  
Coordenação:  
Zenaida de Oliveira Norais Carneiro (UEFS/Fapesb/CNPq)  
Marizete Fagundes de Oliveira Lacerda (UEFS/Fapesb)

Universidade Estadual de Feira de Santana  
Departamento de Letras e Artes  
Núcleo de Estudos de Língua Portuguesa (NELP)  
Projeto Vozes do Sertão em Dados: história, poezia e formação do português brasileiro (DOHS - Documentos Históricos do Sertão)  
Processo CNPq 401433/2009-9, Consepe 102/2009  
Coordenação:  
Zenaida de Oliveira Norais Carneiro (UEFS/Fapesb/CNPq)

Fonte: Site CE-DOHS.

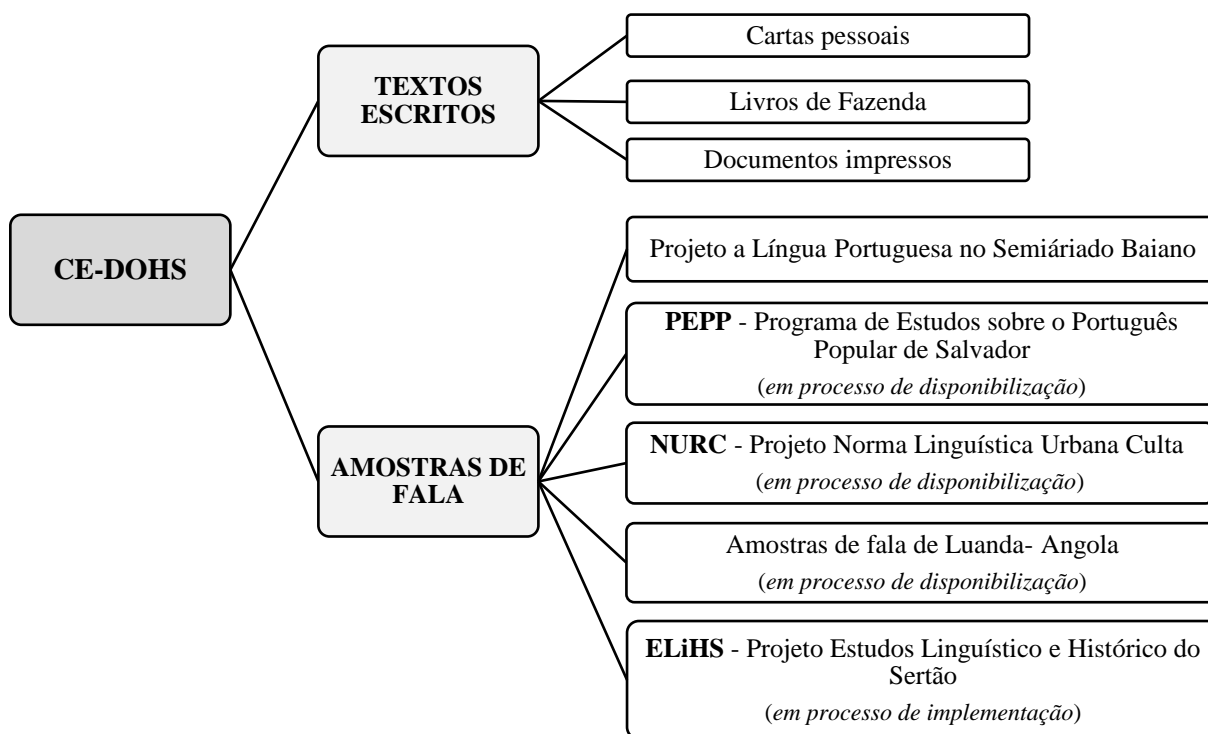
O *corpus* disponibilizado nos permite acesso à edição fac-similar em diferentes formatos e à versão anotada, bem como informações complementares sobre os acervos, tour virtual pelos locais dos acervos (3D VR 360 vídeos); localização geográfica dos autores; sites dos acervos; distribuição do *corpus* por metade de século em número de palavras, dentre outras funcionalidades que envolvem a busca eletrônica.

Nesse contexto, o CE-DOHS tem se configurado como uma plataforma digital de *corpora* do PB, sendo “*composto por documentos que já ultrapassaram as fronteiras dos sertões, representando a maior parte do Brasil*” e tem centrado seus esforços na edição filológica, digital e eletrônica e na anotação morfossintática de textos de acervos manuscritos e orais, tendo em vista o estudo do português brasileiro. O banco de dados, de forma geral, tem como pretensão contribuir com ações voltadas à pesquisa, formação de recursos humanos e transferência de conhecimento para a sociedade, por meio do uso do banco *online*, assim como o ELiHS, o que os torna complementares.

No atendimento às agendas locais e nacionais, o CE-DOHS tem contribuído para os estudos sócio-históricos e linguísticos diante de dois grandes campos – escrito e oral – conforme *Figura 32*.



**Figura 30** – Estrutura geral do Projeto CE-DOHS



Fonte: Elaborada pela autora

A plataforma permite acesso exclusivo ao banco de dados orais, que pode ser acessado através do link de acesso < <http://www5.uefs.br/cedohs/amostrasorais/index.html>>, em que o ELiHS já é citado, como projeto em desenvolvimento (cf. Figura 31).

**Figura 31**– Plataforma eletrônica CE-DOHS: amostras de fala

Fonte: <http://www5.uefs.br/cedohs/amostrasorais/index.html>

A plataforma CE-DOHS passa, atualmente, por uma ampliação do banco de dados<sup>70</sup> e, no que tange às amostras orais, temos até o momento 102 inquéritos disponíveis para *download*, já em sua versão *xml*, conforme *Quadro 17*:

**Quadro 17** – CE-DOHS: amostra oral disponível para *download*

		AMOSTRA/EDIÇÃO FAC-SIMILADA		CORPUS DISPONÍVEL EM XML	DESCRIÇÃO
PROJETO A LÍNGUA PORTUGUESA NO SEMIÁRIO BAIANO	Fase I (1997-1999)	Piemonte da Diamantina/Piabas		12 entrevistas	Este <i>corpus</i> oral dispõe de entrevistas com falantes analfabetos ou pouco escolarizados, de ambos os sexos, divididos em três faixas etárias. A metodologia de coleta foi feita nos moldes da sociolinguística quantitativa (LABOV, 1972, 1982).
		Chapada Diamantina	Barra dos Negros/Bananal	12 entrevistas	
			Mato Grosso	12 entrevistas	
	Fase II (1999-2003)	Nordeste/Jeremoabo	Casinhas	12 entrevistas	
			Tapera	06 entrevistas	
			Lagoa do Inácio	06 entrevistas	
		Paraguaçu	Matinha	12 entrevistas	
			Feira de Santana	-	
PROGRAMA DE ESTUDOS SOBRE O PORTUGUÊS POPULAR DE SALVADOR (PEPP)		Salvador	48 entrevistas	Este <i>corpus</i> foi constituído entre os anos de 1998 e 2000 e coordenado pela professora Dr. <sup>a</sup> Norma da Silva Lopes (UNEB). As gravações foram transcritas segundo as normas definidas pelo Projeto NURC/Salvador. Os informantes estão distribuídos em quatro faixas etárias: de 15 a 24 anos, de 25 a 35 anos, de 45 a 55 anos e de 65 anos em diante. Quanto à sua escolaridade, varia entre mínima, de 1 a 4 anos (pouca escolaridade), e máxima, de 11 anos de permanência na escola (média escolaridade	

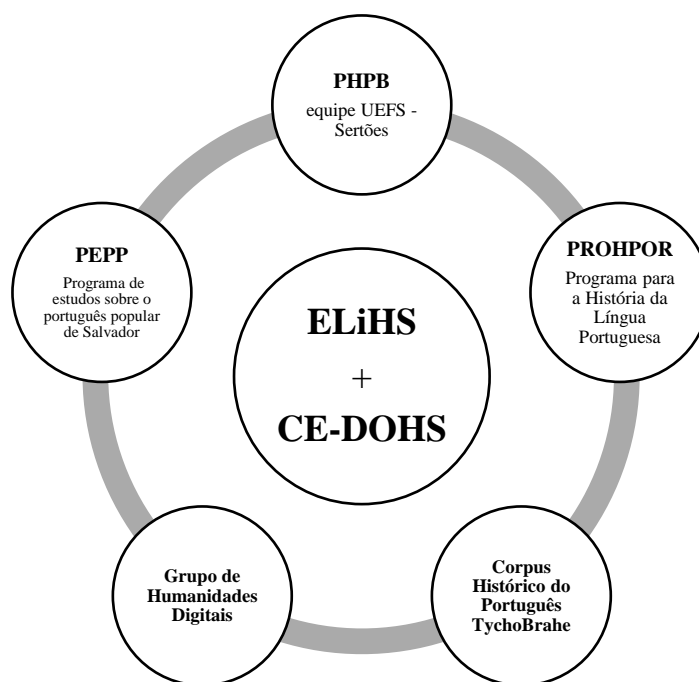
Fonte: Adaptado do *site* do CE-DOHS.

<sup>70</sup>Cabe relatar que o Projeto contará com a publicação de 02 artigos científicos a serem publicados pela Revista Labor, a saber: *Para um estudo da formação do português brasileiro: descrição, representatividade e potencialidades do corpus colonial do CE-DOHS* (Cardoso; Carneiro; Lacerda, no prelo) e *CE-DOHS: um banco de dados sociolinguísticos para a história do português brasileiro* (Santiago; Lacerda; Brito; Carneiro, no prelo).

Na tentativa de atender agendas específicas do CE-DOHS, o Projeto ELiHS retomará a constituição de amostras orais, com objetivo de disponibilizar 48 (quarenta e oito) amostras de língua oral, com falantes analfabetos ou pouco escolarizados (até o quinto ano), de ambos os sexos, divididos em três faixas etárias, de 06 (seis) comunidades afro-brasileiras do Território de Identidade de Irecê-Bahia, contribuindo para o conhecimento da realidade linguística brasileira e, de forma específica, para o estudo da língua falada em áreas do semiárido baiano.

O ELiHS/UNEB, enquanto projeto parceiro do CE-DOHS/UEFS, estabelece uma interinstitucional rede de apoio e parcerias, atendendo a importantes agendas de pesquisa no campo da Linguística (*cf.* Figura 32):

**Figura 32** – ELiHS e CE-DOHS: uma parceria na constituição de *corpora* orais



Fonte: Elaborada pela autora.

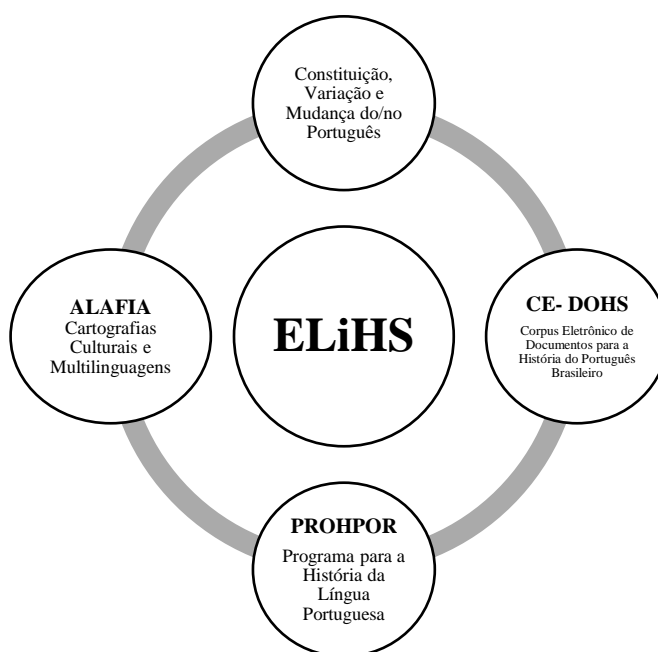
A importante parceria entre o CE-DOHS e o ELiHS retoma a idealização de constituição/ampliação de um banco de dados sociolinguísticos, agora enquadrado em padrões metodológicos, visando a sua implementação no contexto dos metadados<sup>71</sup>, visto que é preciso

<sup>71</sup> Conferir artigo intitulado *CE-DOHS: um banco de dados sociolinguísticos para a história do português brasileiro* (Santiago *et al.*, 2021), em que as autoras apresentam o controle sócio-histórico realizado na constituição do Corpus Eletrônico de Documentos Históricos do Sertão (CE-DOHS). Trata-se de um banco de dados sociolinguísticos, que reúne mais de um milhão de palavras, disponibilizando documentos editados em diferentes versões e seus metadados, além do perfil sociocultural dos escreventes.

pensar que os repositórios representam o legado deixado por nossos Projetos e os avanços tecnológicos têm acompanhado as Ciências Humanas. O CE-DOHS caracteriza-se como uma inspiração na idealização de execução do ELiHS, já que, “considerando as relações sociolinguísticas, isto é, a interação entre o sistema de relações linguísticas e as disposições nas quais ele se atualiza, o CE-DOHS é hoje visto como um dos *corpora* mais significativos” (Santiago *et al*, 2021, p. 325).

Também se faz importante relatar que o Projeto ELiHS se delinea dentro de uma rede de pesquisa já estabelecida (*cf.* Figura 33), participando de quatro diretórios de pesquisas científicas, o que demonstra seu potencial de propagação dos resultados vindouros, incentivando as pesquisas no âmbito da descrição linguística, em especial, de comunidades afro-brasileiras.

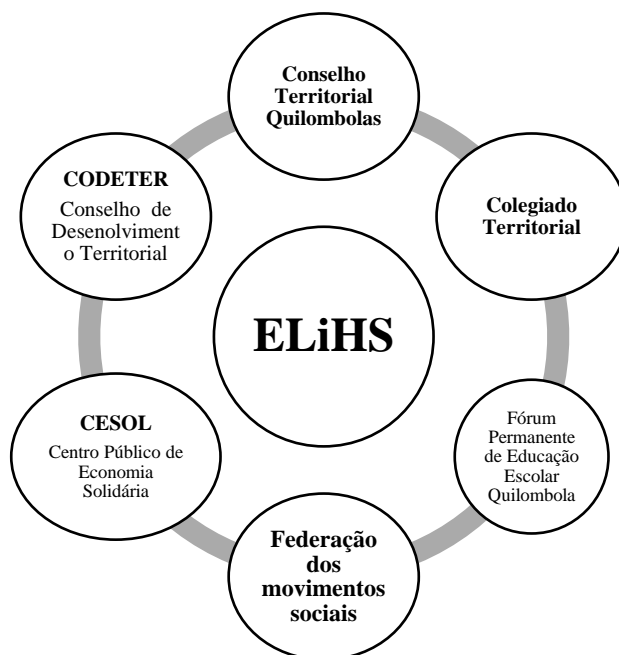
**Figura 33** – Projeto ELiHS: Rede de Diretório de Pesquisas no Brasil



Fonte: Elaborada pela autora.

Somando as parcerias interinstitucionais (*cf.* Figura 32) e de pesquisa (*cf.* Figura 33), o Projeto ELiHS já apresenta, ainda em seu processo de implementação, um importante potencial de impacto regional, com uma rede de parceria já firmada na reunião do CODETER (Conselho de Desenvolvimento Territorial), realizada no dia 27 de julho de 2021.

**Figura 34** – Projeto ELiHS: Rede de parceria Territorial



Fonte: Elaborada pela autora.

Como prerrogativa o Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), todo e qualquer projeto de pesquisa deve constituir-se dentro da preocupação dos impactos não somente científicos, mas também, e principalmente, sociais, privilegiando a qualidade e não a quantidade da produção científica, bem como a relevância e impacto no seu sentido mais amplo. Assim nos reconhecemos enquanto Projeto.

Nesta tese, trazemos como um dos produtos “um Projeto, um *corpus* e uma plataforma *online*”. Dessa forma, a criação e implementação do Projeto ELiHS insere-se entre os projetos que contemplam a pluralidade de comportamentos linguísticos, oferecendo material linguístico para a descrição de variedades ainda não catalogadas, assim como amostras de fala das comunidades afro-brasileiras do Território de Irecê. Sem dúvidas, a proposta de mapeamento do português falado em comunidades afro-brasileiras se delinea pela importância da realização de pesquisas regionais, na tentativa de uma descrição linguística sistematizada.

### **4.3 Conectividade ELiSH: armazenamento e disponibilização do acervo**

O Projeto ELiHS trata-se de um banco de dados digital, acessível na rede mundial de computadores, que reúne *corpora* orais representativos do português popular de comunidades

afro-brasileiras, disponibilizando áudios e transcrições inéditas, na tentativa de realizar diálogos interdisciplinares e interinstitucionais, alcançando outros projetos que objetivam constituir banco de dados orais e contribuir, ainda que indiretamente, para a compreensão do processo de formação do Português Brasileiro, especificamente em sua vertente popular.

Se antes o pesquisador realizava suas pesquisas em laboratórios restritos a grupos de pesquisa, agora, com o advento das novas Tecnologias de Comunicação e Informação (TICs), as possibilidades de interação com outros grupos, de qualquer parte do mundo, se tornam mais efetivas e potencializadas, através da ampla divulgação e da rede mundial de computadores, visto que a Ciência Aberta permite a “(a) publicação completa, franca e rápida dos resultados; (b) ausência de restrições relativas a propriedade intelectual; e (c) transparência, radicalmente aumentada, em fases de pré e pós-publicação de dados, atividades e decisões dentro de grupos de investigação (Maurer, 2003, p. 4 *apud* Jacobetty, 2010, p. 10), como propõe o Projeto ELiHS.

Portanto, esta plataforma online, gratuita, a ser descrita nesta seção, baseia-se na concepção de que

os custos de acesso impostos pelos detentores de direitos de propriedade intelectual sobre o conhecimento técnico e científico têm consequências nefastas para programas de pesquisa exploratória, considerados vitais no paradigma da Economia do Conhecimento. [...] a ‘ideia’ detida por um indivíduo não tem o mesmo valor que assume quando é disseminada comunitariamente, acumulando os benefícios da partilha e aplicação dessa mesma ideia, originando o *spillover* de benefícios.

[...]

A própria dinâmica do progresso científico exige um certo grau de abertura do conhecimento. Só desse modo é possível, no seio das diferentes áreas de pesquisa, validar descobertas e reduzir a duplicação desnecessária de esforços de investigação (Jacobetty, 2010, p. 11-12, grifo do autor).

É neste contexto de acessibilidade e inovação que o Projeto ELiHS divulga seus primeiros resultados, produtos desta tese, numa interrelação com a proposta da Linguística de *Corpus*, visto que ela

[...] se ocupa da coleta e exploração de corpora, ou conjuntos de dados linguísticos [...] que foram coletados criteriosamente com o propósito de servirem para a pesquisa de uma língua ou variedade linguística. Como tal, dedica-se à exploração da linguagem através de evidências empíricas, extraídas por meio de computador (Berber Sardinha, 2000, p. 325).

Assim, iniciaremos com a proposta de armazenamento de *corpora*, seguindo com as possibilidades de exploração, através das TICs; dentro das agendas posteriores de trabalho, temos o objetivo de fazer a etiquetagem morfológica e anotação sintática do *corpus*<sup>72</sup>.

<sup>72</sup> Proposta de organização de um *corpus* anotado morfológica e sintaticamente para a execução de buscas automáticas e/ou semiautomáticas, bem como a recuperação rápida e confiável de grande quantidade de dados para pesquisas linguísticas. Seguiremos o padrão adotado no âmbito do *Corpus* Anotado do Português Histórico Tycho Brahe (tycho.iel.unicamp.br), coordenado por Charlotte Galves (UNICAMP). Para maiores

Com afirma Berber Sardinha (2000), a Linguística de *Corpus* tem experimentado um crescimento vertiginoso nos últimos anos e tem tido um impacto considerável na linguística, portanto o banco de dados coletados, com rigor metodológico (cf. 4.2.2), através do Projeto ELiHS, não poderia deixar de acompanhar esta tendência mundial<sup>73</sup>. Assim, chegamos à conclusão de que a Plataforma *online* do Projeto sustenta-se no tripé tecnológico e científico que envolve (a) as Tecnologias de Comunicação e Informação; (b) a Ciência Aberta; e, por fim, mas não menos importante, (c) a Linguística de *corpus*.

**Figura 35** – Tripé tecnológico e científico da Plataforma do Projeto ELiHS



Fonte: Elaborada pela autora.

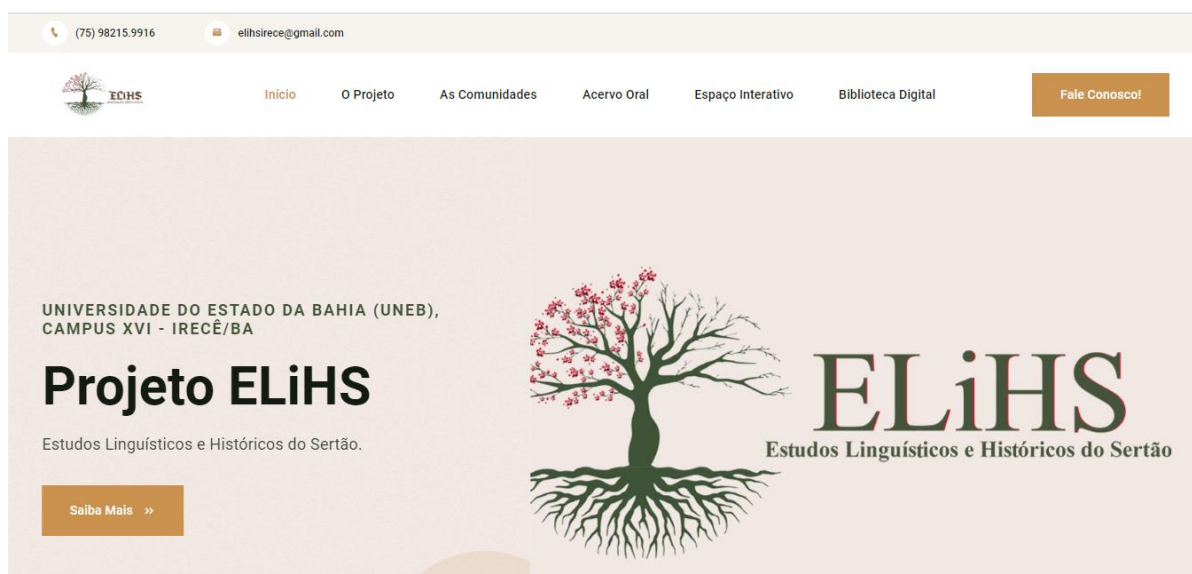
Segue apresentação detalhada da plataforma *online*, a começar pela página inicial<sup>74</sup> (cf. Figura 36).

esclarecimentos sobre anotação sintática, sugerimos a leitura de *Syntactic Annotation System*. Disponível em: <http://www.tycho.iel.unicamp.br/~tycho/corpus/manual/syn-fm.html>.

<sup>73</sup> “Hoje em dia, a Linguística de Corpus é de grande influência na pesquisa lingüística, em vários centros. Na Grã-Bretanha, um dos centros mais desenvolvidos, várias universidades (Birmingham, Brighton, Lancaster, Liverpool, Londres, etc.) dedicam-se à pesquisa baseada em corpus para a descrição dos mais variados aspectos da linguagem. A pesquisa em instituições britânicas tem possibilitado tanto a teorização quanto a criação de corpora e de materiais de apoio em diversas áreas. Igualmente, nos países escandinavos (Noruega, Suécia e Dinamarca) existem centros estabelecidos dedicados à Linguística de Corpus com um papel atuante há vários anos” (Berber Sardinha, 2000, p. 327-328).

<sup>74</sup> Pode ser acessado através do link <https://projetoelihs.com/index.php>.

**Figura 36** – Página inicial da plataforma do Projeto



Fonte: Acervo pessoal.

O *site* foi pensado/desenvolvido<sup>75</sup> dentro de uma configuração moderna, funcional e interativa, por isso sua *home page* é apresentada numa perspectiva dinâmica, em que facilmente o internauta/pesquisador consegue passear pelos *menus*, bem como ter acesso rápido a informações básicas sobre o Projeto – contato, logo, sede, parceria. Na *home*, ainda disponibilizamos 03 (vídeos) com depoimentos que marcam a importância do Projeto, a saber:

- (i) **Dos monitores**<sup>76</sup>, que narram a importância do Projeto ELiHS em suas trajetórias acadêmicas;
- Allane de Oliveira - Monitora do Projeto desde 2022, aluno do 4<sup>a</sup> semestre do curso de Letras, moradora da comunidade Lagoa do Zeca – Canarana (BA);
  - Cauê Oliveira – Monitor do Projeto desde 2022, aluno do 6<sup>a</sup> semestre do curso de Letras, morador de São Gabriel (BA);
  - Cleilton dos Santos - Monitor do Projeto desde 2021, aluno do 6<sup>a</sup> semestre do curso de Letras, morador da Comunidade de Floresta II – Canarana (BA);
  - Jaiane Moreira – Monitora do Projeto desde 2022, aluna do 6<sup>a</sup> semestre do curso de Letras, moradora da comunidade Volta Grande – Lapão (BA);

<sup>75</sup> O site foi estruturado/organizado por Dayane Lemos e monitores do Projeto ELiHS, e desenvolvido por Alan Pinheiro, Tecnólogo em Análise e Desenvolvimento de Sistemas.

<sup>76</sup> As Autorizações de imagem e voz estão arquivadas no acervo físico do Projeto.



- Geisiane Gonçalves – Monitora do Projeto desde 2022, aluno do 3<sup>a</sup> semestre do curso de Pedagogia, moradora da comunidade Olhos D'Aguinha – Boninal (BA);
  - Genilde Medrado - Monitora do Projeto desde 2022, aluna do 8<sup>a</sup> semestre do curso de Letras, moradora da comunidade Aguadas – Barra do Mendes (BA);
  - Guilherme de Oliveira – Monitor do Projeto desde 2023, aluno do 2<sup>a</sup> semestre do curso de Letras, morador da Comunidade Sítio Novo – Jussara (BA);
  - Joanderson Borgres – primeiro monitor do Projeto, desde 2019, aluno do 7<sup>o</sup> semestre de Letras e morador da Comunidade Volta do Angico – Canarana (BA);
  - Haiane Pereira – Monitora do Projeto desde 2019, aluna do 8<sup>a</sup> semestre do curso de Pedagogia, moradora da Comunidade de Canabrava – América Dourada (BA);
  - Marina Barreto – Monitora do Projeto desde 2021, aluna do 6<sup>a</sup> semestre do curso de Letras, moradora de Lapão (BA);
  - Nalanda Pimenta - Monitora do Projeto desde 2021, aluna do 6<sup>a</sup> semestre do curso de Letras, moradora de Canarana (BA);
- (ii) **Da Direção e coordenadores**<sup>77</sup>, que descrevem a importância do Projeto EliHS dentro do tripé ensino, pesquisa e extensão;
- Ana Karine Loula – Diretora do Departamento de Ciências Humanas e Tecnologias (DCHT) da UNEB, *campus* XVI – Irecê;
  - Maria Dorath Sodré – Coordenadora do Curso de Especialização em Educação do Campo;
  - Hilderlândia Machado – Coordenadora do Colegiado de Letras.
- (iii) **Do Programa de Pós-graduação stricto sensu**<sup>78</sup>
- Joabson Figueiredo – Coordenador do Mestrado em Estudos Africanos, Povos Indígenas e Culturas Negras (PPEGAFIN);

---

<sup>77</sup> As Autorizações de imagem e voz estão arquivadas no acervo físico do Projeto.

<sup>78</sup> As Autorizações de imagem e voz estão arquivadas no acervo físico do Projeto.

- (iv) **Da comunidade externa**<sup>79</sup>, que narra os impactos do Projeto ELiHS no Território de Irecê.
- Valdicléia Marques- membro do Conselho Quilombola do Território de Irecê; coordenadora titular do Fórum Permanente de Educação Quilombola; Conselheira Estadual de Desenvolvimento Territorial (CEDETER); membro do Conselho de Desenvolvimento Territorial (CODETER);
  - Lécio Batista – Presidente da associação da Comunidade Lagoa do Martinho
  - Thais Ferreira – Coordenadora do CESOL (Centro Público de Economia Solidária do Território de Irecê).

Os vídeos apresentados estão acessível na página inicial do site e caracterizam-se como um pequeno demonstrativo do impacto do Projeto ELiHS no âmbito social, econômico e cultural das comunidades afro-brasileiras do Território de Irecê, em curto, médio e longo prazos, bem como na formação acadêmica dos monitores, que têm compreendido diariamente a importância da pesquisa científica em seu processo formativo universitário.

Como relatado em narrativas audiovisuais, o trabalho desenvolvido pelo Projeto, desde 2019, tem exercido importante papel não apenas na proposta linguística de interiorização dos estudos sociolinguísticos e sócio-históricos brasileiros, mas também na formação de recursos humanos<sup>80</sup> e no direcionamento de holofotes às comunidades que estão “para além do litoral” – como sugere o título desta tese –, que vivem e resistem no Território de Irecê.

Nesse sentido toda a estrutura de *layout* foi pensada não apenas na comunidade acadêmica, mas também na externa; indivíduos que ansiavam por um espaço representativo para seus povos e comunidades. Assim, buscamos chegar a uma configuração moderna,

---

<sup>79</sup> As Autorizações de imagem e voz estão arquivadas no acervo físico do Projeto.

<sup>80</sup> O Projeto, que conta atualmente com 13 monitores, tem aberto espaço para formação de indivíduos mais conscientes da realidade linguística brasileira e, então, mais qualificados profissionalmente, tendo em vista que estão em um curso de licenciatura em Letras. Nesse sentido, destacamos os subprojetos em desenvolvimento por alguns dos monitores, a saber:

- a. Allane de Oliveira - Processo de escolarização no sertão baiano: uma viagem descritiva a comunidade afro-brasileira de Lagoa do Zeca (BA);
- b. Jaiane Moreira - Novos olhares sobre o existir e resistir na comunidade afro-brasileira de Volta do Angico: um estudo etnofotográfico;
- c. Joanderson Borges –Aspectos culturais e identitários da comunidade quilombola de Volta do Angico Canarana (BA);
- d. Marina Barreto – Os jogos *online* como suporte pedagógico para o Ensino de Língua Portuguesa: um convite à metodologia ativa;

funcional e interativa, em que contamos com *menus* (parte superior da *Figura 34*) a serem descritos no *Quadro 18*:

**Quadro 18** – Descrição funcional do *menu* do *site*

<b>O PROJETO</b>	<b>Descrição Geral</b>		Descrição geral do Projeto, bem como seus objetivos a curto e longo prazos.
	<b>Frentes de Trabalho</b>		Descrição das agendas de trabalho.
	<b>Quem somos</b>		Apresentação da Equipe ELiHS, organizada em coordenações, com acesso ao currículo lattes.
<b>AS COMUNIDADES</b>	<b>Dados sociodemográficos</b>		Mapeamento e caracterização das comunidades afro-brasileiras certificadas de Território de Identidade de Irecê.
	<b>Cidade – Comunidade</b>		Acesso à Galeria fotográfica e descrição sócio-histórica.
<b>ACERVO ORAL</b>	<b>Metodologia</b>	<b>Seleção das comunidades e dos participantes</b>	Descrição da metodologia adotada.
		<b>Gravação</b>	
		<b>Transcrição</b>	
	<b>Acervo</b>	<b>Volta do Angico</b>	Disponibilização dos áudios e das transcrições – formato PDF, <i>xml</i> e modernizada (em processamento).
<b>ESPAÇO INTERATIVO</b>			Jogo <i>online</i> para o ensino da concordância nominal de número.
<b>BIBLIOTECA DIGITAL</b>			Depósito para as pesquisas já desenvolvidas sobre as comunidades afro-brasileiras do Território, bem como as que serão desenvolvidas na execução do Projeto, na tentativa de ampliar a visibilidade dessas produções científicas.
<b>CONTATO</b>			Disponibilização de <i>e-mail</i> e contato telefônico do Projeto.

Fonte: Elaborado pela autora.

O menu *O projeto* e *As comunidades* foram desenvolvidos a partir da proposta desta tese, conjugada com a proposta do Projeto ELiHS, de buscar motivações na sócio-história do português brasileiro, conforme proposições de Rosa Virgínia Mattos e Silva, sintetizadas por Lobo (2015, p. 71), já discutidas na seção 2.

1. A história linguística do Brasil não se restringe à história da língua portuguesa no Brasil, nem à história do português brasileiro.
2. O português brasileiro emerge em contexto multilíngue: o contato linguístico é, pois, elemento constitutivo da sua formação.
3. Na cena linguística do Brasil colonial, destacam-se três atores principais: o português europeu, as línguas gerais indígenas e o português geral brasileiro.

4. Africanos e afrodescendentes foram os principais difusores da língua portuguesa no Brasil e os principais formatadores do português brasileiro em sua variante social majoritária - o português popular brasileiro.
5. O passado sócio-histórico-linguístico do Brasil deverá ser interpretado para a compreensão do português brasileiro «heterogêneo e variável, plural e polarizado»

A descrição sintética apresentada no *site* pode ser retomada nesta tese através da seção 4.2, na tentativa de compreendermos melhor as propostas que impulsionaram, impulsionam e impulsionarão as frentes de trabalho do Projeto ELiHS. No menu *O Projeto*, ainda é possível conhecer toda a equipe que colabora para que esta plataforma esteja acessível na rede mundial de computadores.

No menu *As Comunidades*, além de apresentarmos uma descrição generalizada sobre o contexto sociodemográfico (melhor descrito na seção 4.1.4) das comunidades reconhecidas no Território de Irecê, apresentamos, ineditamente, uma breve descrição de cada comunidade, a partir do questionário aplicado pessoalmente nas visitas, e uma galeria de fotos representativas de cada comunidade<sup>81</sup>.

O menu *Acervo Oral* conta com a descrição metodológica, da seleção de participantes e comunidades ao processo de gravação e transcrição, já descritos nas seções 4.2.2.2, 4.2.2.3 e 4.2.2.4, respectivamente. Neste espaço, os internautas/pesquisadores poderão acessar, gratuitamente, os áudios e transcrições inéditas<sup>82</sup>, com o devido rigor metodológico e sócio-histórico.

Todo o *corpus* do Projeto poderá ser facilmente acessado, através da rede mundial de computadores, e o internauta/pesquisador poderá ter acesso ao áudio, através do *link* que o direcionará ao *drive*<sup>83</sup> do Projeto; é possível solicitar automaticamente o *download* das transcrições, em 03 (três) versões, a saber: PDF (download automático), *xml* (disponibilização *online* em nova aba) e modernizada (em construção<sup>84</sup>), conforme *Figura 37*.

Sem dúvidas, o livre acesso ao conhecimento e ao que se produz no âmbito do Projeto ELiHS é um incentivo à realização de novas pesquisas, nas mais distintas perspectivas teóricas e metodológicas, e uma tendência, tendo em vista a proposta da Ciência Aberta (*Open Science*),

---

<sup>81</sup> Por falta de recursos financeiros e apoio de poderes públicos, só conseguimos realizar visitas às comunidades da cidade de Canarana (09 comunidades visitadas; restam Largo dos Mirandas e Novo Horizonte), Lapão (08 comunidades visitadas, resta Largo), Itaguaçu da Bahia (04 comunidades visitadas), Jussara (02 comunidades visitadas) e Central (02 visitadas), totalizando 25 comunidades com registros fotográficos.

<sup>82</sup> No momento, só está disponibilizado o *corpus* da comunidade Volta do Angico; no entanto, já temos mais 06 transcrições em processo de revisão final, da comunidade Lagoinha, em São Gabriel, e 03 da comunidade de Lagoa do Gaudêncio, em Lapão.

<sup>83</sup> Os áudios serão disponibilizados via *drive*, devido ao tamanho dos arquivos, que implicaria no carregamento da página do site no momento do acesso.

<sup>84</sup> A opção de texto modernizado ainda não está disponível; encontra-se nas próximas agendas de trabalho do Projeto.

um novo modo de fazer ciência, diante da possibilidade dada através das TICs, que potencializam as interações e interdisciplinaridades científicas. Sendo assim, o Projeto ELiHS apoia a proposta da Ciência Aberta, reagindo à privatização do conhecimento (Jacobetty, 2010)<sup>85</sup>.

Não podemos deixar de ressaltar que essa disponibilização dos dados empíricos pelo projeto ELiHS, numa perspectiva do presente para o passado, fundamenta-se também na afirmação da professora Rosa Virgínia Mattos e Silva, na década de 90 do século passado, de que os dados nunca foram ou serão menos importantes do que os quadros teóricos, visto que estes podem mudar, enquanto aqueles permanecem, podendo ser analisados segundo diferentes teorias, em qualquer tempo; acrescentamos a essa afirmação o fato de agora poderem ser acessados em qualquer lugar do mundo, através da rede mundial de computadores – o que tem feito projetos como o *Corpus* Eletrônico de Documentos Históricos do Sertão <[http://www5.uefs.br/cedohs/xml/corpora\\_o.html](http://www5.uefs.br/cedohs/xml/corpora_o.html)>, *Corpus* Anotado do Português Histórico TychoBrahe <[tycho.iel.unicamp.br](http://tycho.iel.unicamp.br)>, dentre outros.<sup>86</sup>

Contamos também com nossa biblioteca digital (em construção), espaço em que depositaremos as pesquisas já desenvolvidas sobre as comunidades afro-brasileiras do Território, bem como as que serão desenvolvidas na execução do Projeto, na tentativa de ampliar a visibilidade dessas produções científicas. Atualmente, o banco de dados já conta com algumas produções, conforme Quadro 19.

**Quadro 19** – Acervo da Biblioteca Digital

Capítulos de livros	SILVA, A. P. da; LEMOS, D. M.. Variável concordância verbal de terceira pessoa do plural: contribuições para o perfil sociolinguístico da comunidade de remanescentes quilombolas de Lagoinha. In: SILVA, J.; LEMOS, D.. <b>Educação, Linguística e Literatura</b> : múltiplas faces do ensino. Tutoiá/MA: Editora Diálogos, 2022.
	ALMEIDA, A. R. D de; LEMOS, D. M.. Variação linguística entre o “nós” e o “a gente” na comunidade quilombola lagoinha, São Gabriel-BA. In: SILVA, J.; LEMOS, D. M. (org.). <b>Estudos Linguísticos</b> : (novos) caminhos. v. 2. Tutoiá/MA: Editora Diálogos, 2021.
Monografias	SILVA, Alex Porfiro da. <b>A variável concordância verbal de terceira pessoa do plural</b> : contribuições para o perfil sociolinguístico da comunidade de remanescentes quilombolas de lagoinha. 2021. Monografia do curso de Letras, Universidade do Estado da Bahia (UNEB), campus XVI, Irecê, 2021.
	SANTOS, Maria Damiana Gonçalves dos. <b>Uma abordagem da variação no uso da preposição <i>ni</i> em corpus oral da comunidade de remanescentes quilombolas Volta do Angico</b> . 2021. Monografia do curso de Letras, Universidade do Estado da Bahia (UNEB), campus XVI, Irecê, 2021.

<sup>85</sup> Conferir Jacobetty (2010), que afirma que essa noção de privatização decorre da relação entre Estado, Ciência e Academia, percebida no fim do século XX. Disponível em: <https://repositorio.iscte-iul.pt/handle/10071/3024>.

<sup>86</sup> Marquilhas e Hendrickx (2016) discorrem sobre *Avanços nas Humanidades Digitais* e citam, entre os projetos de *corpora* históricos reconhecidos internacionalmente, o CE-DOHS (p. 21) e o Tycho Brahe. O capítulo faz parte do livro, *Manual de Linguística Portuguesa*, organizado pelas professoras Ana Maria Martins e Ernestina Carrilho, publicado pela De Gruyter.

	AMORIM, Jeane Souza. <b>Processo de variação na concordância nominal de número na comunidade de remanescentes quilombolas de Lagoinha em São Gabriel-BA.</b> 2021. Monografia do curso de Letras, Universidade do Estado da Bahia (UNEB), campus XVI, Irecê, 2021.
	ALMEIDA, Ana Raquel Dantas de. <b>Variação linguística entre o “nós” e o “a gente” na comunidade quilombola lagoinha, São Gabriel-BA.</b> 2020. Monografia do curso de Letras, Universidade do Estado da Bahia (UNEB), campus XVI, Irecê, 2020.

Fonte: Elaborada pela autora.

Este ainda pequeno repositório se restringe a pesquisas realizadas no âmbito do Projeto ELiHS, através de amostras inéditas representativas do português popular de comunidades afro-brasileiras do Território de Irecê – Volta do Angico e Lagoinha. No entanto, este é um espaço que estará aberto à divulgação das pesquisas das mais diversas perspectivas e áreas de conhecimento, tendo como único critério a abordagem de temáticas que envolvam as comunidades do Território.

Portanto, em nosso espaço, todos terão acesso ao banco de dados, a dados sócio-históricos (ou metadados), à galeria fotográfica e ao acervo digitalizado das produções científicas realizadas sobre as comunidades afro-brasileiras do Território de Irecê, bem como a um espaço interativo para o ensino da concordância nominal de número, através do jogo *online*, concatenando a nova perspectiva assumida dentro da Sociolinguística – Sociolinguística Educacional (Bortoni-Ricardo, 2004) –, numa efetiva relação entre análise linguística e a aplicabilidade ao ensino de Língua Portuguesa, conforme descrição realizada na seção que segue.

#### 4.3.1 *Concordância*: Descrição do jogo *online* como suporte pedagógico para o ensino da concordância nominal número

Diversos fenômenos de variação e mudança linguística, como a concordância nominal de número – objeto de análise desta tese –, têm sido desprestigiados no processo de ensino-aprendizagem de língua portuguesa na Educação Básica, desconsiderando sua recorrência em textos orais e escritos produzidos por estudantes, em seu percurso de aprendizagem. Neste sentido, após os resultados estatísticos prototípicos da Sociolinguística Quantitativa (*cf.* seção 5.3), alcançados através da análise do *corpus* oral representativo do português popular da comunidade afro-brasileira Volta do Angico, passamos a lançar o questionamento em torno da aplicabilidade destes resultados ao ensino de língua portuguesa, dada a tendência assumida a partir da Sociolinguística Educacional (Bortoni-Ricardo 2004).

Nesse sentido, compreendermos que

antes de se delinearem as estratégias didáticas que visam melhorar a eficiência do ensino fundamental no Brasil, é necessário proceder ao levantamento dos traços estigmatizados dos dialetos das classes menos favorecidas. Isso permitirá aos educadores promover um diagnóstico das possíveis interferências dialetais, bem como elaborar métodos e material didático que atendam ao problema de diglossia que se instalou nas escolas (Bortoni-Ricardo, 2005, p. 38).

É considerando essa urgência no campo da ciência das linguagens que passamos a nos questionar de que forma os resultados estatísticos alcançados nesta tese poderiam ser explorados na interface educacional, promovendo a valorização da heterogeneidade peculiar à língua materna e o reconhecimento de construções linguísticas variáveis no Português Brasileiro, no que tange à concordância nominal de número. Esforços que se somam ao trabalho de tantos para uma educação linguística plena e de qualidade. No nosso ponto de vista, a Linguística tem um compromisso necessário com a Educação, havendo a necessidade de uma aplicação didático-pedagógica das pesquisas realizadas, entendendo, como Ilari (2003, p. 111), que um dos principais desafios da Linguística no século XXI é “mudar a cara do ensino da língua materna”<sup>87</sup>.

Assim, surge a proposta do jogo *online*, *ConcordAção*<sup>88</sup>, em que temos os usuários como protagonistas no processo de ensino-aprendizagem e uma relação direta entre a variação na concordância nominal de número e o uso de tecnologias no contexto da sala de aula – uma proposta que chega neste século, para revolucionar as metodologias de ensino tradicionais, com novas perspectivas educacionais.

Trata-se de um grande desafio melhorar a qualidade do ensino de língua materna, e os estudos em torno da ciência da linguagem não poderiam deixar de dedicar-se a isso. Nesse sentido, sabe-se que

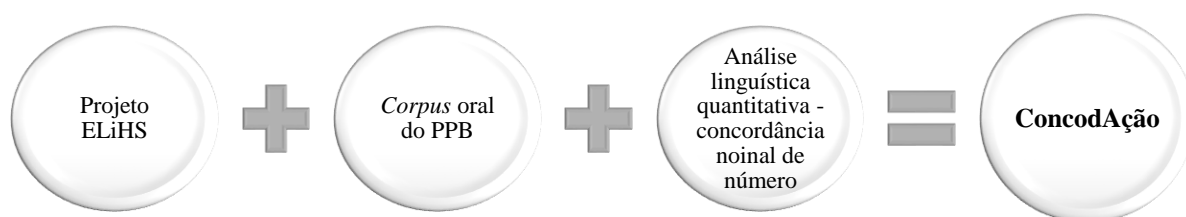
[...] a interface entre a Sociolinguística – especialmente a de cunho variacionista – e o ensino de língua materna tem sido objeto de pesquisadores e de obras de divulgação no cenário nacional (cf. BORTONI-RICARDO, 2004; GORSKI; COELHO, 2006; BAGNO, 2007; 2009, entre outros. *Entretanto, apesar dos avanços, sua aplicação empírica, na sala de aula, é incipiente* (Coan; Freitag, 2011, p. 173-174, grifo nosso).

Neste ponto, de uma aplicação empírica, na sala de aula, surge a proposta do Projeto ELiHS em desenvolver um espaço interativo *online*, voltado ao ensino de Língua Portuguesa e sua aplicabilidade em contextos reais – comunidades afro-brasileiras –, a partir de resultados de pesquisas científicas realizadas no âmbito do Projeto (cf. Figura 36).

<sup>87</sup> Sobre o ensino de língua materna também se debruçou Mattos e Silva (1996; 2004b).

<sup>88</sup> Eu, coordenadora do Projeto ELiHS, e Marina Barreto, bolsista de Iniciação Científica do Projeto, somos responsáveis pela elaboração do aplicativo *ConcordAção*. O resultado apresentado ineditamente nesta tese será apresentado, em forma de comunicação, no *X Encontro das Ciências da Linguagem Aplicadas ao Ensino* (ECLAÉ), um evento organizado pelo *Grupo de Estudos Linguísticos e Literários do Nordeste* (GELNE).

**Figura 36** – Eixo construtor do Jogo *ConcordAção*



Fonte: Elaborada pela autora.

O jogo estará disponível para acesso *online* no *site* do Projeto<sup>89</sup>, no menu *Espaço Interativo*, em que o usuário/estudante poderá facilmente acessar a tela inicial do *ConcordAção* (cf. Figura 37).

**Figura 37** – Tela inicial do jogo interativo



Fonte: Acervo do Projeto.

A tela inicial foi pensada e desenvolvida cuidadosamente na tentativa de trazer elementos que caracterizassem as comunidades do Território – céu ensolarado, habitações simples, plantação e criação de animais –, bem como a presença do Projeto, através da árvore da

<sup>89</sup> Diante de limitações financeiras, tivemos que optar apenas por sua utilização no *site* do Projeto e não, também, em forma de aplicativo, que poderia ser acessado por aparelhos *smartphones*. Tal escolha se justifica pela falta de acesso a *smartphones* e à internet pelos adolescentes das comunidades do Território, sendo esse nosso público alvo. Lembramos que muitas das comunidades aqui mapeadas encontram-se em estado de vulnerabilidade social e, muitas vezes, o acesso à internet se restringe a espaços escolares, em suas salas de informática.



Barriguda<sup>90</sup>. O cenário posto possibilita maior engajamento dos usuários/estudantes, uma vez que estarão imergindo em um contexto que mescla virtualidade e realidade, tornando a proposta do jogo mais atrativa e próxima de seu contexto sociocultural.

O bode e o “monstro cacto” (cf. Figura 37) caracterizam-se como personagens centrais do jogo, por isso aparecem de forma destacada; eles direcionarão a dinâmica do jogo. A próxima tela dará aos usuários/estudantes a opção de 02 (dois) caminhos, a saber:

- (a) Diálogo entre os avatares<sup>91</sup> do Projeto;
- (b) Acesso ao jogo de labirinto<sup>92</sup>.

Na opção (a), teremos a presença de 02 (dois) avatares, auto-representativos para o Projeto ELiHS, a saber:

- (i) Joanderson Borges – Primeiro monitor bolsista do Projeto ELiHS, desde 2019, morador da comunidade Volta do Angico – Canarana (BA);
- (ii) e Haiane Soares – Primeira monitora voluntária do Projeto ELiHS, desde 2019, moradora da comunidade Canabrava – América Dourada (BA).

**Figura 38** – Avatares do Projeto



Fonte: Acervo do Projeto.

<sup>90</sup> Além de simbolizar o Projeto ELiHS, em visita de campo, pudemos observar que a Barriguda está presente em quase todas as Comunidades, ou em suas estradas de acesso.

<sup>91</sup> Os Avatares têm sido caracterizados como uma tendência no mundo virtual dos *Games*, uma forma de convidar os usuários e tornar a proposta do jogo mais atrativa.

<sup>92</sup> Dada a limitação de tempo, estamos apresentando apenas o jogo do labirinto; no entanto, posteriormente, teremos outras opções para os usuários. A proposta é que os monitores de iniciação científica sigam com a elaboração de novas estratégias de jogos *online*, bem como com a abordagem de outros fenômenos representativos do processo de contato linguístico (cf. Lucchesi; Baxter; Ribeiro, 2009).

Os avatares são uma singela homenagem a quem sempre esteve presente de forma tão participativa no Projeto ELiHS e representa, de forma geral, a ancestralidade do povo afro-brasileiro. O diálogo produzido por eles traz uma breve reflexão sobre a diversidade linguística no PB, bem como sobre as normas básicas que regem o uso da concordância de número nos sintagmas nominais, em contextos de formalidade, finalizando com um convite de acesso ao jogo *ConcordAção*.

Assim, o jogo foi pensado, acreditando em que “é tempo de conferir prioridade ao estudo da ‘linguagem das maiorias’” (Bortoni-Ricardo, 2005, p. 37). Nesse sentido, elaboramos estratégias metodológicas que pudessem atender ao papel central da diversidade linguística no ensino de Língua Portuguesa, com base em resultados estatísticos alcançados na análise linguística desta tese, que apontam que a extensão dos constituintes flexionáveis condiciona a não marcação de plural nos SNs, bem como os itens (-) salientes/regulares (*cf.* seção 5.3).

**Quadro 20** – Estratégia para desenvolvimento das perguntas do jogo

ETAPAS DO JOGO	ESTRATÉGIA LINGUÍSTICA DE APLICABILIDADE	
1 <sup>a</sup>	SNs com 02 constituintes flexionáveis	(-) saliente
		(+) saliente
2 <sup>a</sup>	SNs com 03 constituintes flexionáveis	(-) saliente
		(+) saliente
3 <sup>a</sup>	SNs com 04 constituintes flexionáveis	(-) saliente
		(+) saliente
		(+) saliente

Fonte: Elaborada pela autora.

Posta a estratégia, a partir dos resultados alcançados nesta tese, foi pensada a forma que o jogo assumiria, sendo definido o labirinto, por acreditarmos em que ele estimula a concentração, a coordenação motora, bem como o raciocínio lógico/linguístico.

**Figura 39** – Layout do jogo de labirinto do *ConcordAção*



Fonte: Acervo do Projeto.

A natureza competitiva empenhada durante o uso do jogo é um incentivo para refletirem sobre o fenômeno e criarem estratégias para cumprirem as etapas que seguem, colaborando diretamente no processo de ensino-aprendizagem, em relação ao fenômeno da variação na concordância nominal.

Este é o primeiro passo do Projeto ELiHS em direção a uma relação mais efetiva entre a ciência da linguagem e uma educação linguística combativa frente ao preconceito linguístico, afinal acreditamos veementemente

[...] ser um grande equívoco a afirmação de que a variação linguística não deve ser matéria de ensino na escola básica. Assim, a questão crucial [...] é saber como tratá-la pedagogicamente, ou seja, como desenvolver uma pedagogia da variação linguística no sistema escolar de uma sociedade que ainda não reconheceu sua complexa cara linguística e, como resultado da profunda divisão socioeconômica que caracterizou historicamente sua formação (uma sociedade que foi, por trezentos anos, escravocrata), ainda discrimina fortemente pela língua os grupos socioeconômicos que recebem as menores parcelas da renda nacional (Faraco; Zilles, 2015, p. 8).

Que esta proposta inicial se torne um convite à novas propostas metodológicas e que possamos, não tardiamente, estudá-los a partir dos pressupostos que fundamentam a Metodologia da Linguística Ativa<sup>93</sup>, desenvolvida pela professora Eloisa Pilati (2017), que,

<sup>93</sup>“A contribuição educacional dessa abordagem para o ensino de gramática é a sistematização do pensamento gramatical de forma a possibilitar que as diferentes regras gramaticais sejam compreendidas a partir de princípios mais gerais. Essa forma de se conceber a organização gramatical das línguas naturais explicita caráter sistemático da organização linguística e busca promover a compreensão da organização gramatical por meio desses princípios e não por meio da apresentação de diferentes regras. Com isso, é possível promover uma base conceitual para que as aulas de gramática em laboratórios em que fenômenos gramaticais são investigados,

dentre outros objetivos, busca promover a aprendizagem ativa por meio do desenvolvimento de habilidades metacognitivas, acreditando em que “[...] o ensino de gramática, da forma como vem sendo praticado, não tem contribuído nem para a compreensão dos fenômenos gramaticais nem para a formação de cidadãos confiantes em seu saber gramatical tácito e em sua capacidade de expressão linguística” (Pilati, 2017, p. 15). Que Projetos como os desenvolvidos por ela – Gramaticoteca e Gamatikê<sup>94</sup> – inspirem outros tantos projetos que têm focado suas análises na sociolinguística quantitativa, com resultados tão expressivos para a compreensão do processo de formação do PB, mas que, agora, precisam seguir uma nova tendência, numa perspectiva sincrônica, sobre os processos de variação linguística e o ensino de Língua Portuguesa.

Assim, finalizamos esta seção, destacando que a intenção do jogo *ConcordAção* é contribuir, de forma singela, para a didatização da linguística, fazendo chegar às escolas básicas materiais com dados reais de língua, para a abordagem da variação linguística, recomendada por documentos institucionais – Parâmetros curriculares Nacionais (Brasil, 1998) e Base Nacional Comum Curricular (BNCC) (Brasil, 2016); uma contribuição particularmente no que diz respeito à concordância nominal de número. Não temos aqui a pretensão de apontar soluções definitivas para o ensino do fenômeno, mas apresentar possibilidades para um processo de ensino-aprendizagem mais aproximativo, identitário e lúdico e inspirar outros trabalhos no campo da Sociolinguística Educacional<sup>95</sup>, em uma rede de pesquisa e produção que visa não apenas ao impacto científico, mas aos impactos educacional e social.

---

compreendidos de forma profunda, sistematizados e manipulados em contextos diversos por meio de elementos visuais, coloridos e lúdicos” (Pilati; Veneziano, 2023, p. 8962).

<sup>94</sup> “Considerando os pressupostos da Aprendizagem Linguística Ativa, sob o ponto de vista do ensino das virtualidades do sistema gramatical do português, os principais objetivos do aplicativo são contribuir para que os estudantes desenvolvam habilidades relacionadas ao reconhecimento do ‘sistema linguístico’ da Língua portuguesa, sobre os temas aliadas à promoção de experiências lúdicas de aprendizagem” (Pilati; Veneziano, 2023, p. 8962).

<sup>95</sup> Foi, entre outros, também no NELP, do qual faz parte o projeto CE-DOHS, com que o ELiSH mantém parceria, que buscamos inspiração para nosso produto educacional, aqui apresentado. Como parte do projeto NELP da Sala de Aula: diálogos entre ensino, pesquisa e extensão, criado pelas professoras Mariana Fagundes de Oliveira Lacerda e Zenaide de Oliveira Novais Carneiro e atualmente coordenado pela primeira professora, há alguns produtos educacionais disponíveis, como, por exemplo, o *blog Falaê PB* – com um Caderno de Aprendizagens de 80 páginas (<https://falaepb.wordpress.com>) –, desenvolvido como parte da dissertação de Silva (2021), orientada pela professora Mariana Fagundes de Oliveira Lacerda no Mestrado Profissional em Letras (PROFLETRAS/UEFS), e o audiovisual *Tour Virtual pelo Banco CE-DOHS: Coleção Documental Cartas para Vários Destinatários* (<https://www.youtube.com/@nelpnucleodeestudosemlingu8083/playlists>), trabalho desenvolvido, sob a orientação da professora Mariana Fagundes de Oliveira Lacerda, pela estudante de Graduação Mirian Marques Galindo, vinculada ao Programa de Iniciação Científica da UEFS, com bolsa da FAPESB.

## 5 A VARIAÇÃO NA REALIZAÇÃO DE MARCAS DE PLURAL NO SINTAGMA NOMINAL E SUA IMPORTÂNCIA PARA OS ESTUDOS SOBRE A FORMAÇÃO E A DESCRIÇÃO DO PORTUGUÊS BRASILEIRO

“Eu prefiro ser  
Essa metamorfose ambulante  
Do que ter aquela velha opinião  
Formada sobre tudo.”

(Metamorfose ambulante - Raul Seixas)

Nesta seção abordaremos o fenômeno da *variação na concordância de número no sintagma nominal*, apresentando uma breve síntese sobre a ocorrência da variável no português brasileiro. De forma concomitante, faremos um levantamento de trabalhos já realizados acerca da variável investigada, a fim favorecer uma visão mais abrangente sobre o fenômeno, bem como atestaremos resultados que possam contribuir para as análises a serem descritas na *seção 5.3*, quando possível estabelecendo análises comparativas.

As descrições e análises sobre a variação na realização de marcas de plural no sintagma nominal foram realizadas em *corpus* oral da comunidade afro-brasileira Volta do Angico, por meio de uma amostra representativa do português popular, através da perspectiva teórica da Variação Linguística Laboviana ou Sociolinguística Variacionista Quantitativa (Labov, 2008, [1972]), já descrita na *seção 3.2.1*. Antecedendo às análises, teceremos considerações acerca da abordagem metodológica utilizada no tratamento dos dados, bem como a descrição das variáveis predictoras<sup>96</sup> que regem a sistematicidade do fenômeno proposto à análise. As escolhas metodológicas seguem a hipótese da existência de um sistema gerenciador da variação da concordância de número entre os elementos do SN que permite prevermos circunstâncias linguísticas e sociais favoráveis à realização da variante padrão de concordância.

Por fim, apresentaremos as análises apoiando-nos na modelagem estatística de regressão logística com resposta binária, a linguagem de programação denominada R (R CORE TEAM, 2020), notadamente por meio de uma de suas interfaces, o *RStudio* (EQUIPE RSTUDIO, 2020), que reporta quantitativamente os condicionadores da escolha da variante de plural padrão nos

---

<sup>96</sup> Nesta seção de análise, por optamos pelo uso da linguagem de programação R (R CORE TEAM, 2020), utilizaremos as terminologias indicadas por Oushiro (2022, p. 259, grifo nosso), visto que “em análises multivariadas, a variável dependente y é chamada de *variável resposta*, e as variáveis independentes  $x^1$ ,  $x^2$  etc. são chamadas de *variáveis predictoras*. [...] os termos ‘variável dependente’ e ‘variável independente’ não são adequados para modelos de regressão.”

elementos flexionáveis do sintagma nominal. A partir dos resultados buscamos alcançar dois objetivos, a saber: (i) observar se a comunidade tradicional afro-brasileira Volta do Angico registra padrões morfossintáticos de concordância comuns a outras comunidades tradicionais, a exemplo de Helvécia (Andrade, 2003), com vistas a contribuir para os estudos da sócio-história do português brasileiro; e (ii) utilizar os resultados no desenvolvimento de material pedagógico, jogo eletrônico<sup>97</sup>, para o ensino da concordância nominal de número, através da metodologia da aprendizagem linguística ativa (Pilati, 2017), um dos produtos desenvolvidos no âmbito desta tese, através do Projeto ELiHS, na tentativa de colaborar, ainda que timidamente, para o ensino de língua portuguesa.

A importância da análise diante do fenômeno da concordância nominal de número recai sobre a constância no processo de variação – como imprime nossa epígrafe, uma verdadeira *metamorfose ambulante* –, rompendo com padrões rigidamente tradicionais – *e não aquela velha opinião formada sobre tudo* –, e cabe a nós, linguistas, apreciá-la, analisá-la, descrevê-la, para assim repensarmos novas práticas funcionais para o ensino de língua portuguesa, principalmente em escolas **do** e **no** campo.

#### **4.1 Revisitando os princípios gerais da concordância nominal: da prescrição à descrição gramatical**

A variação na concordância no português falado no Brasil tem sido intensamente documentada e estudada no âmbito da sociolinguística, oferecendo, desse modo, a oportunidade de evidenciar as perdas morfológicas tão típicas do português brasileiro. No contexto linguístico de concordância em língua portuguesa é possível citarmos 03 (três) mecanismos gramaticais, quais sejam:

- (i) Gênero – especificamente nominal (**o** amigo / **a** amiga);
- (ii) Pessoa – especificamente verbal (**nós** conversamos / **eles** conversam);
- (iii) Número – Nominal ou verbal.

Dentre os mecanismos citados, destacamos a concordância de número, que, por sua vez, pode ser nominal ou verbal. A verbal ocorre entre o sintagma nominal (doravante SN) sujeito

---

<sup>97</sup> O jogo ficará disponível na plataforma do Projeto. Direcionaremos um capítulo adicional a Tese para apresentação das suas funcionalidades e contribuições para o ensino da Concordância Nominal de Número.

e o verbo (ele **estuda** / eles **estudam**), a nominal ocorre entre os elementos flexionáveis do SN, como em (1), ou entre o SN sujeito e o predicativo, como em (2):

(i) Os elementos constituintes do Sintagma Nominal:

(1) “**As** **peessoas** não tem coragem mais de saí em casa.”<sup>98</sup> (J. P. S.)



(ii) O Sintagma Nominal sujeito e o predicativo do sujeito:

(2) “**As** **peessoas** são **bonitas**.” (J.A.A.S.)



Em (1) e (2) é possível afirmarmos que a concordância nominal de número ocorre a partir da inserção da marca de plural *-s* em todos os elementos flexionáveis do sintagma nominal. Portanto, a marca formal de plural nos nomes em português é representada categoricamente pelo *-s*, considerando que a forma regular de se fazer este tipo de plural no PB é através da adição de um segmento fricativo não labial, cujas realizações podem ser [s], [z], [ʃ], [ʒ] ou [h], estas realizações ocorrem em função de contextos linguísticos específicos ou em função de características regionais dos falantes do português do Brasil (Scherre, 1996, p. 87).

No entanto, a variação no uso das regras de concordância apresenta uma gradação em função das diversas variedades do português brasileiro, que graças aos avanços dos estudos dialetológicos e sociolinguísticos têm sido possível a sistematização descritiva do fenômeno, contrapondo as imposições da tradição gramatical ao apresentar argumentos que comprovam que no campo da oralidade o fenômeno da concordância de número no SN parece não ser tão obrigatório e categórico, especificamente quando se trata dos falantes do PB. Portanto, no PB contemporâneo, a concordância nominal de número dentro do sintagma nominal pode apresentar-se através de 03 (três) possibilidades, a saber:

(i) Inserção da marca de plural em todos os elementos do SN

(3) “Assim **os** **alunos** assim às vez quando chove eles perde aula.” (J. A. S.)

(ii) Marcação de alguns elementos do SN

---

<sup>98</sup> Todas orações apresentadas nesta pesquisa, a título de exemplificação, são retiradas do *corpus* da comunidade afro-brasileira Volta do Angico.

(4) “Cada qual tem os seus empregoØ<sup>99</sup>.” (J. Q.)

(iii) Marcação de plural em um único elemento do SN

(5) “Se saísse pá roça numa chasse as coisaØ prontaØ apanhava.” (M. L.)

Isso implica dizer que, ao se tratar do PB, a concordância nominal de número não é tão invariável como afirmam algumas gramáticas, não é uma regra categórica, mas sim variável conforme certos determinantes linguísticos e sociais (Scherre, 1994). Assim, ao se tratar da variação na concordância de número no SN, podemos afirmar a existência de duas variantes interligadas ao fenômeno: a variante explícita (marcada) e a variante zero de plural (não marcada), como observamos em (6) e (7):

(i) Variante explícita:

(6) “As festas aqui é boa.” (F. P. B.)

(ii) Variante zero de plural:

(7) “Ele cava as covaØ.” (M. L.)

Questões como a concordância (nominal e verbal) têm sido discutidas sob a perspectiva da funcionalidade das marcas linguísticas, no entanto, a partir dos exemplos (6) e (7), percebemos que no português do Brasil é possível eliminar algumas marcas de concordância dentro do SN, respeitando o princípio da economia, preservando as marcas necessárias e eliminando as redundantes, como em (7); ou, de modo contrário, preservando as marcas redundantes, como em (6).

Em geral, a concordância é compreendida como uma relação de solidariedade entre os elementos de um dado enunciado, com a inserção da marca de plural -s em todos os elementos flexionáveis do sintagma. Para Dubois *et al.* (1973, p. 136), concordância é

o fenômeno sintático pelo qual um substantivo ou um pronome pode exercer pressão de alteração formal sobre os pronomes que o representam, os verbos de que ele é sujeito, e os adjetivos ou participios que a ele se referem.

Dentre os muitos gramáticos temos Evanildo Bechara, autor de várias das principais gramáticas da língua portuguesa, que tenta sintetizar a concordância de número no SN, como:

<sup>99</sup> A simbologia Ø utilizada neste trabalho reflete a confluência formal entre o zero plural e a morfologia do singular, caracterizada pelo morfema zero, como define Scherre e Naro (2006, p. 107).



Diz-se *concordância nominal* a que se verifica em gênero e número entre o adjetivo e o pronome (adjetivo), o artigo, o numeral ou o particípio (palavras determinantes) e o substantivo ou pronome (palavras determinadas) a que se referem (Bechara, 2009, p. 543, grifo do autor).

Após a definição essencialmente gramatical para *concordância nominal*, Bechara (2009, p. 544-554) aponta uma série de contextos de ordem formal e semântica em que vocábulos determinantes se subordinam aos seus núcleos, mediante a ação da concordância. Os contextos são os seguintes:

#### A. CONCORDÂNCIA DE PALAVRA PARA PALAVRA

- (i) *Há uma só palavra determinada* – a palavra determinante irá para o gênero e número da palavra determinada;
- (ii) *Há mais de uma palavra determinada*:
  - Se as palavras determinadas forem do mesmo gênero, a palavra determinante para o plural e para o gênero comum, ou poderá concordar, principalmente se vier anteposta, em gênero e número com a mais próxima;
  - Se as palavras determinadas forem de gêneros diferentes, a palavra determinante irá para o plural masculino ou concordará em gênero e número com a mais próxima.
- (iii) *Há uma só palavra determinada e mais de uma determinante* – a palavra determinada irá para o plural ou ficará no singular, sendo, neste último caso, facultativa a repetição do artigo.

B. CONCORDÂNCIA DE PALAVRA PELO SENTIDO – a palavra determinante pode deixar de concordar em gênero e número com a *forma* da palavra determinada para levar em consideração, apenas, o *sentido* em que esta se aplica;

#### C. OUTROS CASOS DE CONCORDÂNCIA NOMINAL

Até o momento parece estar tudo muito bem estruturado aos olhos da gramática tradicional, mas cabem algumas contestações a respeito da eficácia e viabilidade dessas conceituações/regras apresentadas, por muitos gramáticos, em especial pelo autor da *Moderna Gramática Portuguesa*, de que nos servimos para uma reflexão. Vejamos:

- (i) Qual o real sentido do estabelecimento de regras, aparentemente fixas, e, em seguida, de forma bastante tímida, apresentar as observações ou as tradicionais e constantes *exceções*?
- (ii) Por que afirmar ao leitor a existência de três contextos referentes à *concordância nominal*, sendo que um deles (*Outros casos de concordância*) aglomera 21 (vinte e um) contextos específicos?
- (iii) Será que realmente é possível categorizar um fenômeno tão variável quanto a *concordância nominal de número* ou as gramáticas caminharão para o “mundo sem fim” das exceções do português do Brasil?

A primeira reflexão surge das muitas exceções apresentadas pelas gramáticas tradicionais. Na *Moderna Gramática Portuguesa* podemos observar que após conceituar a *concordância de palavra para palavra*, em que há uma só palavra *determinada*, o autor traz a seguinte informação:

OBSERVAÇÃO: Os nomes femininos como *sentinela*, *guarda*, *guia* e assemelhados, quando aplicados a pessoas do sexo masculino, mantêm o gênero feminino, e levam para este gênero os determinantes a eles referidos: a *sentinela avançada* (Bechara, 2009, p. 545, grifo do autor).

Quando se trata da *concordância de palavra para palavra*, em que há *mais de uma palavra determinada* o autor apresenta mais três “observações”, uma delas diz: “Se as palavras determinadas se referirem a uma só pessoa ou coisa, impõe-se o singular do determinante” (Bechara, 2009, p. 545). Essas “observações” vão se fazendo presentes em boa parte dos contextos de ordem formal e semântica apresentados pelo autor, porém cabe ressaltar que o fenômeno das exceções gramaticais não se restringe a essa gramática específica, mas sim a toda e qualquer gramática que tente tratar o português como homogêneo.

A segunda reflexão nos faz entender que a intenção de aglomerar 21 (vinte e um) contextos específicos em apenas um, faz com que se evite defini-las como EXEÇÕES, uma vez que iria transparecer ao leitor que, no caso da concordância nominal, há mais exceções do que regras.

A terceira e última reflexão incorre no grande debate existente entre os tradicionalistas e os sociolinguistas, sobre a variabilidade inerente a língua falada, particularmente, ao se tratar do português falado no Brasil.

No que tange a visão formalista temos Câmara Jr. (2000), linguista que desenvolveu muitos trabalhos e inseriu-se na corrente da linguística estruturalista. Nessa perspectiva, dita formal, a língua é analisada fora de seu contexto de uso, considerando-se apenas aspectos internos. No que tange a concordância acredita-se que essa ocorre mediante uma “coerção formal”, entendendo que esse é um fenômeno exclusivamente estrutural. Câmara Jr. (2000, p. 31) acredita que ao se tratar da concordância entende-se que “[...] na flexão há obrigatoriedade e sistematização coerente. Ela é imposta pela própria natureza da frase”.

A variação no uso do plural é abordada pelo autor, mas apenas para alertar sobre a realização do arquifonema /S/, o qual se atualiza, na oralidade, através das fricativas não-labiais ou sibilantes, /s/ - /z/, ou chiantes /s'/ - /z'/, atualizações essas não representadas na grafia do fonema indicador de plural (-s), pontuando que tais variantes estão relacionadas à região do falante e são, sobretudo, alomorfias fonologicamente condicionadas.

Diante de uma superficial amostragem das concepções formalistas, a noção de concordância tem a ver com a concepção de língua como objeto invariável. Atribui-se, então, à estrutura da língua toda responsabilidade pelo seu funcionamento.

Enquanto os estudos formalistas baseiam-se na estrutura da língua, a gramática normativa baseia-se no registro escrito – o literário, cuja função é o exercício estético da linguagem – e, portanto, não diz respeito a padrões reais de uso. Desse modo, tais formas de se explicar o fenômeno da concordância nominal de número são pouco elucidativas no que se refere ao seu emprego enquanto prática social, visto que o português vernacular do Brasil apresenta variações nos processos de concordância de número, o que o distancia da visão de homogeneidade e de rigidez pregadas pelas gramáticas tradicionais que tomam como base o português padrão de Portugal.

Essa visão tradicional tem favorecido a propagação da ideia que a variante explícita de plural é a variante de prestígio, enquanto a variante zero (não marcada) é imprópria, ou simplesmente errada e restrita a grupos específicos, em especial os não escolarizados. Porém, diante das pesquisas realizadas no campo sociolinguístico, já é incontestável que a concordância de número se realiza variavelmente tanto nos dialetos populares, quanto na chamada norma culta, no Brasil<sup>100</sup>, caracterizando-se, então, como um fenômeno da oralidade sem exclusões de grupos minoritários e/ou majoritários.

É neste contexto heterogêneo, multivariável e plural que a variação na concordância nominal de número torna-se tema de discussão quando se trata da gênese do português do Brasil, tanto por aqueles que veem suas características como o resultado de um processo secular de deriva, apenas intensificado pelo contato do português com as línguas indígenas e africanas (Naro; Scherre, 2007), quanto por aqueles que relacionam os processos de variação que se observam hoje no português do Brasil, sobretudo em suas variedades populares, ao amplo e massivo contato entre línguas que marca a história sociolinguística do país (Baxter; Lucchesi, 1997; Lucchesi, 2003, 2008).

À vista disso, pensar na variação da concordância de número no SN é entender que o português falado no Brasil apresenta características peculiares ao ser comparado com o português europeu, características essas que incitam pesquisas como esta que trabalha com *corpus* popular representativo de comunidade rural afro-brasileira. Só assim começaremos a trilhar o caminho que nos conduzirá a resultados que nos apontem e ratifiquem que a morfologia nominal do PB é *uma metamorfose ambulante* – português vernacular – e não aquela velha

---

<sup>100</sup> Conferir *Quadro 21*.

*opinião formada sobre tudo* – gramática tradicional – e que este quadro é assegurado pelas condições sócio-históricas em que se formou a língua portuguesa no Brasil, tornando a variação na concordância nominal um dos efeitos mais notáveis da realidade linguística bipolarizada do PB.

### 5.1.1 A revisão bibliográfica: contribuições de alguns estudos

No que tange as análises que envolvem a variação da concordância de número no sintagma nominal, essas têm início com o estudo desenvolvido por Jeroslow, em 1974, que abarca uma pesquisa na área rural cearense, e faz uma descrição estruturalista dos diversos subsistemas linguísticos do falar dessa área e, dentre os diversos aspectos descritos, registra a concordância de número verbal e nominal, mas esses estudos ainda não tinham o aporte teórico metodológico laboviano. Sendo assim, destacam-se como pioneiros os estudos realizados por Braga e Scherre, em 1976, que analisaram dados de 07(sete) falantes do Rio de Janeiro, de classe social e procedência geográfica distintas; o de Braga (1977) que desenvolve sua dissertação com dados de 07 (sete) falantes da classe média e baixa do Triângulo Mineiro, com três níveis de escolaridade; o de Lemle e Naro (1977) que, utilizando os princípios labovianos da teoria da variação, inicia a pesquisa sobre concordância verbo/sujeito com 20 (vinte) falantes semi-escolarizados da zona urbana Rio de Janeiro; o de Scherre (1978) que analisou, em sua dissertação de mestrado, dados de 10 (dez) falantes da zona urbana Rio de Janeiro, com diferentes níveis de escolaridade.

Em 1988, Scherre retoma o tema e “reanalisa” a concordância de número na sua tese de doutorado, intitulada *Reanálise da concordância nominal em português*<sup>101</sup>, hoje considerada como uma obra de referência ao se tratar do tema, já que através desse trabalho a autora analisa todos os mecanismos que regem o fenômeno da variação da concordância no SN no português falado no Rio de Janeiro, analisando-o de forma criteriosa nas duas abordagens de análise, a atomística/mórfica e a sintagmática/não-mórfica. Em suas considerações finais Scherre (1988, p. 277-278) apresenta os fatores que condicionam ou não a variação da concordância no SN, nas perspectivas de análise supracitadas. A autora afirma que seus dados mostraram exercer influência das seguintes variáveis:

---

<sup>101</sup> Analisa a concordância nominal no PB com dados extraídos do *corpus* Censo do Programa de Estudos sobre o Uso da Língua (PEUL), grupo de pesquisa vinculado ao Departamento de Linguística e Filologia da Faculdade de Letras da Universidade Federal do Rio de Janeiro – UFRJ.

- **Na perspectiva atomística:** Marcas precedentes e posição; Saliência fônica: dimensões processos e tonicidade; Relação dos elementos não nucleares em função do núcleo e posição dos elementos nucleares no SN; Formalidade dos substantivos e adjetivos; Grau dos substantivos e adjetivos; Animacidade dos substantivos; Contexto fonético/fonológico seguinte; e Função resumitiva.
- **Na perspectiva não-atomística:** Pluralidade do contexto; Configuração sintagmática do SN; Saliência fônica; Grau de formalidade do SN; Pluralidade do SN; Função textual; e Localização do SN.

A tese de doutorado realizada por Scherre (1988) serviu e serve como base e/ou apoio bibliográfico para pesquisas posteriores que trataram e/ou tratam da concordância nominal. Nessa sessão, também precisamos dar destaque à tese de Lopes (2001)<sup>102</sup>, intitulada *Concordância nominal, contexto linguístico e sociedade*, na qual a autora apresenta resultados da análise sobre a variação da concordância de número no SN nas falas popular e culta na cidade de Salvador, Bahia, em quatro diferentes faixas etárias, 03 (três) graus de escolarização, nos 02 (dois) sexos, e em 02 (dois) diferentes grupos étnicos, identificando fatores linguísticos e sociais que mais condicionam a sua variação. Lopes (2001) também trabalhou com amostras de fala do português popular, a partir do banco de dados do Programa de Estudo do Português Popular de Salvador (PEPP)<sup>103</sup>.

Diferentemente de Scherre (1988), Lopes (2001) trabalhou apenas na perspectiva atomística, apresentando também os fatores que condicionam o uso padrão, marcado, para a concordância de número no SN. A pesquisadora (2001, p. 368-369) chega às seguintes conclusões:

- (i) São mais alvo de concordância elementos com mais material fônico na oposição singular/plural;
- (ii) Elementos à esquerda do núcleo recebem mais marca de plural que elementos nucleares;

---

<sup>102</sup> A análise foi pautada em inquéritos com indivíduos usuários do que se convencionou chamar de Norma Urbana Culta – NURC –, ou seja, com sujeitos portadores de diploma de curso superior.

<sup>103</sup> O PEPP, coordenado pela professora doutora Norma da Silva Lopes (UNEB), é formado por 48 (quarenta e oito) gravações de entrevista de aproximadamente quarenta minutos, com a seguinte caracterização: os informantes estão distribuídos em quatro faixas etárias: de 15 a 24 anos, de 25 a 35 anos, de 45 a 55 anos e de 65 anos em diante. A escolaridade dos informantes dessa amostra varia entre mínima, de 1 a 4 anos (pouca escolaridade), e máxima, de 11 anos de permanência na escola (média escolaridade). As amostras já estão disponíveis em linguagem xml., na plataforma do CE-DOHS, através do link [http://www5.uefs.br/cedohs/xml/corpora\\_o.html](http://www5.uefs.br/cedohs/xml/corpora_o.html).

- (iii) Elementos à direita do núcleo têm menos probabilidade de receberem marca de plural;
- (iv) Contextos antecedentes com numeral promovem mais concordância subsequente em segunda posição que contextos com marca formal antecedente;
- (v) Elementos em 3<sup>a</sup>, 4<sup>a</sup> ou 5<sup>a</sup> posição tendem a fazer mais concordância quando são antecedidos por mais de uma marca do que quando antecedidos por zero.
- (vi) Elementos são mais alvo de plural quando finalizam a sentença;
- (vii) Elementos com contexto subsequente com vogal sonora têm mais probabilidade de não fazerem concordância que contexto com vogal surda.

Na tentativa de conhecermos um pouco mais sobre a produção científica do Brasil, no que tange à morfologia nominal de número do Português Brasil, em especial em sua modalidade oral, elaboramos o *Quadro 21*, em que apresentamos alguns estudos sobre o fenômeno da variação na concordância nominal. Os trabalhos registrados são reflexo de um breve levantamento na Plataforma de Teses e Dissertações da Capes<sup>104</sup> através da palavra-chave de busca “concordância nominal”, entre os anos de 2000 à 2022. O quantitativo, não exaustivo, pretende demonstrar que tal variação não se restringe a variedades regionais, mas que apresenta ocorrências documentadas nos mais diversos espaços brasileiros, evidenciando a hipótese de que tais fatos são mais de natureza diastrática do que diatópica. Vejamos:

---

<sup>104</sup> Disponível em: <https://catalogodeteses.capes.gov.br/catalogo-teses/#/>.

**Quadro 21** – Panorama dos estudos sobre variação na concordância nominal desenvolvidos nas últimas décadas no Brasil (2000 – 2022)<sup>105</sup>

REGIÃO	ESTADO	DISSERTAÇÕES	TESES
NORTE	Amapá		RIBEIRO, C. M. da R. <i>Contato Linguístico e a Concordância de Número no Sintagma Nominal no português de Oiapoque/AP</i> . 2018. Doutorado em Linguística Instituição de Ensino: Universidade Federal do Rio De Janeiro, Rio de Janeiro.
	Amazonas		MARTINS, F. S. <i>Variação na concordância nominal de número na fala dos habitantes do Alto Solimões (Amazonas)</i> . 2013. Doutorado em Linguística Instituição de Ensino: Universidade Federal de Santa Catarina, Florianópolis.
	Pará	FORO, F. M. R. <i>A marcação de número no sintagma nominal em falantes de Belém</i> . 2005. Mestrado em Letras: Linguística E Teoria Literária Instituição de Ensino: Universidade Federal do Pará, Belém.	
		SANTOS, M. L. F. dos. <i>A concordância de número no sintagma nominal na fala de belenenses</i> . 2003. Mestrado em Letras: Linguística e Teoria Literária Instituição de Ensino: Universidade Federal do Pará, Belém.	
NORDESTE	Alagoas	MARQUES, A. K. L. de O. <i>A concordância de número no sintagma nominal: uma análise sociolinguística da língua falada em Maceió</i> . 2016. Mestrado em Linguística e Literatura Instituição de Ensino: Universidade Federal de Alagoas, Maceió.	
		SILVA, F. G. Da. <i>Alagoanos em São Paulo e a Concordância Nominal de Número</i> . 2014. Mestrado em Linguística Instituição de Ensino: Universidade de São Paulo, São Paulo.	MARQUES, A. K. L. de O. <i>Concordância nominal: um olhar sociolinguístico sobre a língua falada em Alagoas</i> . 2021. Doutorado em Linguística e Literatura Instituição de Ensino: Universidade Federal de Alagoas.
	Bahia	MEIRA, G. A. <i>Estudo comparativo entre as normas popular e culta do português de Vitória da Conquista: concordância nominal de número</i> . 2015. Mestrado em Linguística Instituição de Ensino: Universidade Estadual do Sudoeste Da Bahia, Vitória da Conquista.	LOPES, N. da S. <i>Concordância nominal, contexto linguístico e sociedade</i> . 2001. Doutorado em Letras e Linguística Instituição de Ensino: Universidade Federal da Bahia, Salvador.

<sup>105</sup> As referências expostas no *Quadro 21* e não se configuram como base de discussão no corpo da tese não serão replicadas na lista final de Referências, conforme solicita a ABNT.

		GUIMARAES, M. A. de S. <i>Variação na concordância nominal de número no Português popular de Vitória da Conquista – BA: contribuições para compreensão da sócio-história do português do Brasil</i> . 2014. Mestrado em Linguística Instituição de Ensino: Universidade Estadual do Sudoeste da Bahia, Vitória da Conquista.	
		DORIA, T. P. L. <i>Um estudo da variação no sintagma nominal: a concordância de número em comunidades baianas</i> . 2014. Mestrado em Estudos Linguísticos Instituição de Ensino: Universidade Estadual de Feira de Santana, Feira de Santana.	
		LEMOS, D. M. <i>Português brasileiro e português angolano: variação na concordância nominal de número</i> . 2014. Mestrado em Estudos Linguísticos Instituição de Ensino: Universidade Estadual de Feira de Santana, Feira de Santana.	
		ANDRADE, P. R. de. <i>Um fragmento da constituição sócio-histórica do português do Brasil: variação na concordância nominal de número em um dialeto afro-brasileiro</i> . 2003. Mestrado em Letras e Linguística Instituição de Ensino: Universidade Federal Da Bahia, Salvador.	
	<b>Ceará</b>	TABOSA, M. V. S. <i>A variação na concordância nominal de número no falar do Cariri cearense</i> . 2016. Mestrado em Linguística Instituição de Ensino: Universidade Federal do Ceará, Fortaleza.	
	<b>Maranhão</b>	OLIVEIRA, M. A. F. de. <i>Os usos linguísticos e seus reflexos sociais: uma análise Sociolinguística da estrutura do sintagma nominal, em sentenças de Partes do Dia, da São Luís oitocentista</i> . 2021. Mestrado em Letras Instituição de Ensino: Universidade Federal do Maranhão, São Luís.	
		TEIXEIRA, V. P. <i>Variação linguística e fluxos migratórios: a concordância nominal de número na fala dos moradores do bairro Campo de Belém do município de Caxias – MA</i> . 2017. Mestrado em Letras Instituição de Ensino: Fundação Universidade Federal do Piauí, Teresina.	
	<b>Pernambuco</b>	SILVA, C. K. B. da. <i>Variação da concordância nominal em produção oral e escrita de alunos do ensino fundamental e médio de Belo Jardim-PE: assimetria entre fala e escrita?</i> . 2017. Mestrado em Letras Instituição de Ensino: Universidade Federal de Pernambuco, Recife.	



<b>CENTRO-OESTE</b>	<b>Brasília</b>	DIAS, M. C. A. C. <i>A variação na concordância nominal: um contraste entre o urbano e o rural na fala brasiliense</i> . 1993. Mestrado em Linguística Instituição de Ensino: Universidade de Brasília, Brasília.	
	<b>Goiás</b>	PEREIRA, C. C. <i>O apagamento do plural em sintagmas nominais numa comunidade de fala da cidade de Goiás</i> . 2008. Mestrado em Linguística Instituição de Ensino: Universidade Federal de Uberlândia, Uberlândia Biblioteca.	
<b>SUL</b>	<b>Paraná</b>	FONSECA, T. da L. W. da. <i>Concordância nominal em Guarapuava, Paraná</i> . 2016. Mestrado em Letras Instituição de Ensino: Universidade Estadual do Centro-Oeste, Guarapuava.	
	<b>Rio Grande do Sul</b>	OLSEN, M. J. <i>Concordância nominal no sintagma nominal nas áreas urbana e rural de Caxias do Sul: uma análise sociolinguística da fala do ítalo-brasileiro</i> . 2005. Mestrado em Letras e Cultura Regional Instituição de Ensino: Universidade de Caxias do Sul, Caxias do Sul.	SCHNEIDER, S. D. <i>Concordância nominal de número na fala de crianças de 3 a 6 anos de idade do município de Novo Hamburgo: variação linguística na infância</i> . 2012. Doutorado em Letras Instituição de Ensino: Universidade Federal do Rio Grande do Sul, Porto Alegre.
		KOELLING, S. B. <i>A concordância nominal em Porto Alegre (RS): Análise variacionista</i> . 2004. Mestrado em Linguística e Letras Instituição de Ensino: Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre.	MANGABEIRA, A. B. de Azevedo. <i>Variação na concordância nominal, prática social e identidade entre jovens e adultos no Centro do Trabalho (Porto Alegre – RS)</i> . 2016. Doutorado em Letras Instituição de Ensino: Universidade Federal do Rio Grande Do Sul, Porto Alegre.
<b>SUDESTE</b>	<b>Espírito Santo</b>	SCARDUA, J. R. <i>Análise da concordância nominal na fala de Vitória/ES: o linguístico, o social e o estilístico</i> . 2018. Mestrado em Linguística Instituição de Ensino: Universidade Federal do Espírito Santo, Vitória.	SCARDUA, J. R. <i>Variação e estilo: uma investigação da concordância nominal em diferentes situações comunicativas</i> . 2022. Doutorado em Linguística Instituição de Ensino: Universidade Federal do Espírito Santo, Vitória.
		LOPES, L. DE O. J. <i>A concordância nominal de número no português falado na zona rural de santa Leopoldina/ES</i> . 2014. 200 f. Mestrado em Linguística Instituição de Ensino: Universidade Federal do Espírito Santo, Vitória.	LOPES, L. de O. J. <i>“As história são boa. As pessoas, maravilhosa”</i> : análise da variação da concordância verbal e nominal na zona rural de Santa Leopoldina/ES. 2020 382 f. Doutorado em Linguística Instituição de Ensino: Universidade Federal do Espírito Santo, Vitória.
	<b>Minas Gerais</b>	SANTOS, W. N. dos. <i>A concordância nominal de número no português falado em Montes Claros - MG: a influência dos nomes gerais em uma análise variacionista</i> . 2021. 156 f. Mestrado em Estudos Linguísticos Instituição de Ensino: Universidade Federal de Minas Gerais, Belo Horizonte.	

		SANTOS, L. S. M. <i>Sobre a Ausência de Concordância Nominal no Português falado em Pedro Leopoldo-MG: uma Abordagem Variacionista</i> . 2010. Mestrado em Estudos Linguísticos Instituição de Ensino: Universidade Federal de Minas Gerais, Belo Horizonte.	
		RIBEIRO, P. R. O. <i>O perfil sociolinguístico do município de Oliveira Fortes-MG: a concordância nominal e verbal</i> . 2013. Mestrado em Linguística Instituição de Ensino: Universidade Federal de Juiz de Fora, Juiz de Fora.	
		PINHEIRO, L. R. <i>A concordância nominal no português de Belo Horizonte</i> . 2012. Mestrado em LETRAS Instituição de Ensino: Pontifícia Universidade Católica de Minas Gerais, Belo Horizonte.	
	<b>São Paulo</b>	FIAMENGUI, A. H. R. <i>A marcação de pluralidade no SN na fala e na escrita de adolescentes da região de São José do Rio Preto</i> . 2011. Mestrado em Estudos Linguísticos Instituição de Ensino: Universidade Est. Paulista Júlio de Mesquita Filho. São José do Rio.	BUENO, E. S. da S. <i>A regra de concordância de número no sintagma nominal do português popular falado por algumas comunidades de trabalhadores rurais: cortadores de cana no estado de São Paulo</i> . 2001. Doutorado em Letras Instituição de Ensino: Universidade Est. Paulista Júlio de Mesquita Filho, São Paulo.

Fonte: Elaborado pela autora.

Essas e outras tantas pesquisas desenvolvidas no Brasil evidenciam

[...] que o fenômeno da variação na concordância de número no português falado do Brasil, longe de ser restrito a uma região ou classe social específica, é característico de toda a comunidade de fala brasileira, apresentando diferenças mais de grau do que de princípio, ou seja, as diferenças são mais relativas à quantidade de marcas de plural e não aos contextos lingüísticos nos quais a variação ocorre (Scherre; Naro, 1993, p. 38).

Ainda em relação ao *Quadro 21* é notório que na categoria de tese temos apenas o estudo de Andrade (2003) que trata do uso variável da concordância nominal de número em dialeto rural afro-brasileiro, tornando-o dessa forma uma referência para análise contrastiva proposta nesta tese. Posto isto, o destaque principal será dado à essa tese, intitulada *Um fragmento da constituição sócio-histórica do português do Brasil: variação na concordância nominal de número em um dialeto afro-brasileiro*, também considerada uma obra de referência por tratar do uso variável da concordância nominal de número, no dialeto rural afro-brasileiro da comunidade de Helvécia-Ba<sup>106</sup>, buscando explicações da variação na sócio-história do português do Brasil, objetivo similar ao proposto nesta análise.

O panorama exposto até o momento, no âmbito do português do Brasil, demonstra que muitos estudos já foram realizados sobre a concordância nominal, muitos no domínio da língua oral e, ainda, poucos na modalidade escrita<sup>107</sup>, no entanto, este cenário não é tão vantajoso ao se tratar de outras variedades do português europeu, em especial do português falado em países da África.

<sup>106</sup> Em janeiro de 1994, diante do fato de não haver encontrado características mais radicais nas comunidades visitadas até então, Alan Baxter e Dante Lucchesi organizaram uma recolha na comunidade de Helvécia, no Extremo Sul do Estado da Bahia, contando com a valiosa colaboração do antropólogo Marcos Luciano Lopes Messeder, que já havia contribuído com a recolha nas comunidades de Rio de Contas, e com a participação da estudante de Letras da UFBA Nara Barreto. A comunidade de Helvécia tinha um significado especial, porque nela a dialetóloga Carlota Ferreira havia identificado, no início da década de 1960, o que chamou de “remanescentes de um falar crioulo”. (Lucchesi; Baxter; Ribeiro, 2009, p. 16).

<sup>107</sup> Cabe uma breve ressalva para o trabalho intitulado *Concordância nominal: cenas da variação em palcos do século XIX*, publicado por Oliveira, Soledade e Santos (2009), que tratam do fenômeno em uma perspectiva sincrônica, como expõem os autores, de eras idas, passadas. Para tanto utilizam-se de atas escritas por africanos<sup>107</sup>, ao longo do século XIX, analisando o fenômeno através do aparato, na medida do possível, metodológico da Sociolinguística Quantitativa. No artigo chegam a seguinte conclusão que os textos dos africanos, escritos ao longo de duas décadas do século XIX, “[...] se mostraram uma preciosa fonte a servir de base empírica para mostrar que, nos oitocentos, a variação no âmbito da concordância nominal já acontecia e, pelo que tudo indica, de forma sistemática, já que, levando para o passado alguns fatores que regem a concordância no presente, houve consonância com o que se observa nos estudos da concordância nominal em tempos que correm. (Oliveira; Soledade; Santos, 2009, p. 314).

Os autores (2009, p. 277) afirmam em seu artigo não ser este o primeiro trabalho que utiliza como *corpus* a linguagem escrita, já que em 2008, Maria Tereza Borges da Costa, escreve um texto em que procura esclarecer o fenômeno na escrita de alunos do ensino fundamental e médio residentes no município de Ribeira do Pombal – Ba<sup>107</sup>. Os trabalhos de Oliveira, Soledade e Santos (2009) e o de Costa (2008 *apud* Oliveira; Soledade; Santos, 2009, p. 279) têm demonstrado que a língua nas modalidades oral e escrita, mesmo sendo a segunda mais monitorada, apresentam resultados semelhantes no que se refere aos fatores que favorecem/condicionam a variação na concordância de número no sintagma nominal.

Poucos e relativamente recentes são os estudos sobre o fenômeno concordância nominal no variedades africanas do português, de forma ainda preliminar, destacam-se os estudos quantitativos desenvolvidos na variedade de São Tomé (Baxter, 2004; Figueiredo, 2008), na variedade de Cabo Verde (Jon-And, 2009) e os estudos qualitativos nas variedades de Moçambique (Companhia, 2001; Stroud; Perpetua, 1997) e de Angola (Inverno 2004, 2005, 2009; Marques, 1985). Ressaltamos que os trabalhos apontados não dão conta de todos os trabalhos já desenvolvidos na África sobre variação na concordância nominal, mas representam uma evidência de que esse é um fenômeno que merece destaque ao se tratar do português africano, um convite à novas pesquisas, em especial aos que buscam análises contrastivas entre PB e PA.

Dentre os trabalhos mencionados, destacaremos o de Inverno (2004), intitulado *Português vernáculo do Brasil e português vernáculo de Angola: reestruturação parcial vs. mudança linguística*, que, como afirma a autora, constituiu um estudo piloto para uma comparação mais detalhada dos fatores sociais e linguísticos na origem das duas variedades vernáculas do português, a do Brasil e a de Angola, com o objetivo de procurar entender até que ponto os processos sociais e linguísticos que levaram à formação de cada um dos vernáculos terão sido semelhantes ou diferentes.

Inverno (2004, p. 08), em suas análises qualitativas, chega à conclusão que os fenômenos de concordância variável analisados, embora presentes nas duas variedades devido à influência do substrato banto, apresentam, contudo, uma menor estabilidade/cristalização no português vernacular angolano, parecendo, assim, apontar para uma formação mais recente desta variedade.

Temos também o trabalho de Inverno (2009)<sup>108</sup>, intitulado *A transição de Angola para o português vernáculo: estudo morfossintático do sintagma nominal*, que apresenta uma análise dos fatores sociolinguísticos e históricos relevantes para a compreensão da gênese e desenvolvimento do português vernacular de Angola, seguida de uma análise morfossintática do sintagma nominal. A autora formula a hipótese segundo a qual

[...] apenas na segunda metade do século XX um segmento suficientemente significativo da população angolana viria a ter acesso frequente à língua portuguesa e motivação suficiente para aprendê-la, iniciando assim um processo de reestruturação do português, ainda em curso, especialmente no interior entre falantes bilingues do português e línguas banto, via mudança de língua. Esta hipótese é corroborada não só

<sup>108</sup> Os dados linguísticos analisados provêm de um *corpus* oral semi-espontâneo recolhido em 2004, no Dundo, província da Luanda Norte. O *corpus* foi recolhido através de trinta entrevistas informais, falantes que têm o português como segunda língua. Desses informantes, dezesseis apresentam idades compreendidas entre os 5 e os 17 anos e um grau médio de escolarização, 11 falantes apresentam idades compreendidas entre os 20 e os 56 anos e têm pouca ou nenhuma escolarização.

pelo evoluir da situação sociolinguística em Angola nos últimos cinco séculos, mas também pelas características morfo-sintáticas do SN no PVA, as quais denotam um índice de introdução de características de substrato típico de um processo de mudança de língua [...] (Inverno, 2009, p. 11).

Tratando de outras variedades do português, destacamos dois trabalhos de Brandão (2011) que focaliza a concordância de número no sintagma nominal nas variedades urbanas do Português falado no Brasil e em São Tomé, com o objetivo de determinar os fatores que concorrem para a coexistência dos padrões variantes nelas encontrados. Os resultados, advindos de análise fundamentada nos pressupostos da Sociolinguística Variacionista, demonstram que há significativas convergências entre as duas variedades, entre elas as fortes restrições de ordem estrutural e social que norteiam o cancelamento do morfema de número. Damos destaque também ao trabalho de Jon-And (2010), em que são analisadas, de forma quantitativa, algumas variáveis sociais e linguísticas que regem a concordância variável de número no sintagma nominal no português L2 falado em Maputo, Moçambique.

Portanto, de forma geral, esses são alguns dos estudos já realizados sobre concordância nominal sobre o português falado na África, um convite à novas pesquisas que trabalhem com essas variedades do português, tornando possível delinear as convergências e divergências entre as diferentes variedades do português, em especial as variedades faladas em comunidades rurais afro-brasileiras, tomando bom base comparativa dos dados de Helvécia (Andrade, 2003) e Volta do Angico (análise proposta nesta tese), pelo banco de dados da CAPES, até o momento, únicas teses que trabalham o fenômeno da concordância nominal, em *corpus* oral de comunidades rurais afro-brasileiras .

As pesquisas apontadas nesta seção e outras – não revisadas neste levantamento, porém de fundamental importância no campo dos estudos sociolinguísticos já realizados sobre concordância nominal – nos permitem delinear os fatores estruturais que condicionam a marca canônica de plural nos sintagmas nominais e que o mito de homogeneidade linguística não se sustenta empiricamente, ideia apoiada por diversos linguistas. No entanto, o tema da concordância de número plural no SN, ao mesmo tempo que é familiar no âmbito das discussões, em especial do PB, imprime, na atualidade, duas perspectivas em relação a origem dessa variação: a linha contatista e a linha derivista, discutidas teoricamente na *seção 2.3* e agora retomadas na tentativa de contribuir com a hipótese do papel dos contatos linguísticos e os desdobramentos sincrônicos dessa variável no processo de constituição do PB.

Como vimos, o fenômeno da variação na concordância nominal tem sido o foco de muitas discussões no campo linguístico, mais especificamente no campo da sociolinguística, com a realização de trabalhos como o de Scherre e Naro, que hoje são considerados referências

ao se tratar do assunto. As pesquisas atuais têm focado suas análises na busca por motivações linguísticas e sociais que colaborem para sistematização de tal variação e no favorecimento de evidências em torno da polêmica da formação do português brasileiro.

Ao se tratar da formação do PB (*cf.* seção 2.3), duas correntes dividem os centros das discussões: a linha contatista e a linha derivista, as quais se encontram no âmbito das contendas sobre a origem do PB, desde o século XIX. A primeira linha, derivista, acredita que as variações do português brasileiro sejam reflexos do português europeu, e que estas já estariam previstas no sistema linguístico. Scherre e Naro (1998), mesmo com seus poucos dados, o que torna suas conclusões contestáveis, afirmam ter encontrado no português europeu formas variáveis na concordância verbal e nominal, o que faz com que defendam que a variação não é uma marca peculiar ao português do Brasil, mas que além de ocorrer em Portugal, sempre aconteceu em toda a história da língua portuguesa. Estudos nessa vertente colaboram para a hipótese de um processo de deriva natural da língua e descarta a hipótese de ocorrência de processos de crioulização na história da língua portuguesa falada no Brasil.

Lucchesi (1998, 2001, 2008, 2009), Baxter e Lucchesi (1997) acreditam que no Brasil não se formou um crioulo em termos gerais, mas um crioulo leve, no entanto, a língua portuguesa que hoje se fala aqui é herança de um massivo contato linguístico, que veio a gerar uma transmissão linguística irregular (*cf.* seção 2.3.3). No que tange à hipótese do contato linguístico, entende-se que as variedades orais do português popular falado na Bahia oferecem campos de trabalho bastante promissores para o conhecimento da configuração sócio-histórica do PB, na grande área do semiárido baiano, dado o fato de esse ter sido gestado sob o efeito de amplo contato linguístico, durante o período Colonial e ter sido alvo também de um contato mais recente com variedades padrão e não-padrão do PB, propiciado por migrações, a partir da década de 40 do século XX.

Considerando o estudo sob uma perspectiva sócio-histórica, seria de supor que as características atuais da fala popular brasileira resultam de mudanças produzidas no contexto de contato do português com línguas de origem indígena e africana, hipótese assumida por diversos pesquisadores (Lucchesi, 2009; Mattos e Silva, 2001; entre outros). Há ainda que se considerar outros fatos, como por exemplo, as mudanças ocorridas no português lusitano, alvo de mudança durante o período em que o PB foi gestado, o que torna a situação mais complexa, abrindo possibilidades comparativas com variedades atuais do português europeu (PE) contemporâneo.

No que concerne às variedades africanas do português, a questão da variação na concordância nominal de número é geralmente explicada através do contato linguístico, com

possíveis influências das línguas bantu (Baxter 2004, 2009; Figueiredo, 2008; Marques, 1985; Inverno, 2005), mas Jon-And (2010, p. 45) afirma que “por enquanto pode-se constatar que não podemos confiar unicamente na estrutura de línguas bantu como explicação para o favorecimento da marcação de plural da posição adjacente ao núcleo em variedades do português”.

Assim, defenderemos que as variedades orais do português do interior da Bahia, especificamente das comunidade afro-brasileiras do Território de Irecê, podem ser descendentes do chamado *português geral*, que, segundo Mattos e Silva (2001, p. 298-299) teria sido adquirido na oralidade e em situações de aquisição imperfeita, e que, na atualidade, há um processo de convergência da língua popular e da língua padrão/culta, conforme assumido por Lucchesi (2009) para o PB, motivado por contato em grandes centros urbanos e por vetores de standardização linguística, entre os quais a escolarização pode ser a causa principal (Carneiro; Almeida, 2006, 2007).

## 5.2 Metodologia e instrumentos para levantamento e análise dos dados

O material analisado foi extraído do *corpus* do projeto *Estudos Linguísticos e Históricos do Sertão* (ELiHS) (*cf.* seção 4.2), amostra representativa do Português Popular do Brasil de comunidades afro-brasileiras situadas no Território de Identidade de Irecê. Importante relatar que o fato das amostras serem voltadas ao português popular justifica-se pela proposta, assumida nesta pesquisa, de que, embora as variedades cultas e populares do PB se oponha ao PE, o português popular tem uma origem no processo de transmissão linguística irregular (*cf.* seção 2.3.3), diferente do que ocorreu, a princípio, com a variante culta.

As amostras orais serão consideradas como amostras de língua-E (exteriorizadas), portanto o ponto a ser tomado aqui será a assunção de que a mudança se dá durante o processo de aquisição de línguas e que o ambiente de contato linguístico assume um papel importante para o nascimento de uma variedade de língua, como é o caso do português afro-brasileiro, marcado por uma história de contato, de maior ou menor influência, sobretudo em sua vertente popular.

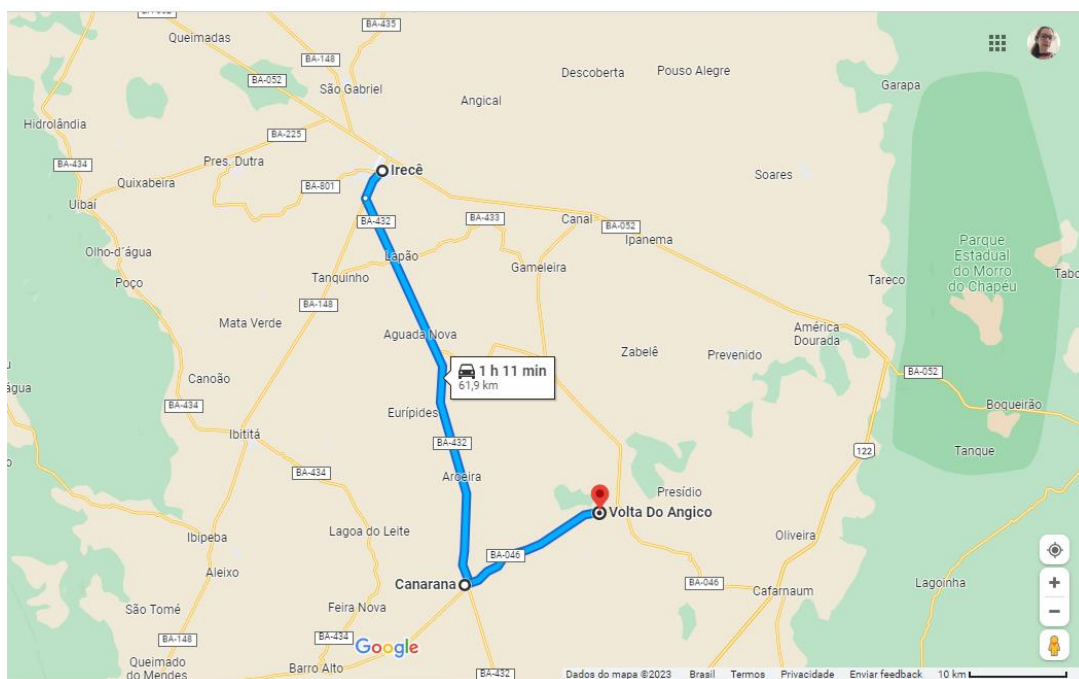
Neste contexto faz-se importante a descrição de características sociodemográficas da comunidade de fala, *corpus* desta pesquisa, com o propósito de apresentar informações que subsidiem a análise sociolinguística da concordância nominal de número no português falado na comunidade afro-brasileira Volta do Angico, pautando-nos na ideia de que as pesquisas sociolinguísticas devem se arrolar não apenas em informações linguísticas, mas, de

forma basilar, em informações sobre o contexto sócio-histórico, cultural e geográfico, os quais, muitas vezes, permitem releituras de dados quantitativos, afinal “o falante encerra em si a síntese do geográfico com o social” (Cardoso, 2010, p. 12).

### 5.2.1 *Lócus* da amostra: comunidade Volta do Angico

A comunidade afro-brasileira Volta do Angico localiza-se à 16,2km de distância da sede, Canarana<sup>109</sup>, e faz divisa ao leste com a comunidade de Mato Verde, ao oeste, com a comunidade de Largo do Miranda, ao sul, com a comunidade de Lagoa do Zeca, e ao norte, faz limite com o Rio Jacaré. Em relação a cidade de Irecê, encontra-se à 61,9km de distância (cf. Figura 40).

**Figura 40** – Mapa de localização da comunidade Volta do Angico – Canarana – Irecê



Fonte: Google maps<sup>110</sup>.

Volta do Angico é reconhecida enquanto Remanescente Quilombola através do processo de nº 01420.010021/2011-16, emitido pela Fundação Cultural Palmares (FCP). Já o processo de regularização territorial através Instituto Nacional de Colonização e Reforma Agrária (INCRA), sob o processo de nº 54160.001937/2014-01, foi iniciado no ano de 2014.

<sup>109</sup> Reconhecida como uma das 20 cidades pertencentes ao Território de Identidade de Irecê.

<sup>110</sup> Disponível em: <https://encurtador.com.br/dwOY7>. Acessado em: 14 jun 2023.



Segundo o presidente da Associação Comunitária Quilombola de Volta do Angico (ASCQVA) a Comunidade tem em média 200 anos de formação, uma das mais antigas do Território, sendo essa formação marcada por ancestrais familiares. O presidente também informa que o nome Volta do Angico é dado em homenagem a uma enorme árvore de angico que ficava localizada no centro da comunidade, há muito tempo atrás<sup>111</sup>.

A Comunidade, e região, é marcada principalmente pelas frequentes secas que podem ser caracterizadas pela ausência, escassez, alta variabilidade espacial e temporal das chuvas. Essas características do meio ambiente condicionam a comunidade a sobreviver principalmente de atividades econômicas ligadas à agricultura e a pecuária. Estas são realizadas de modo a buscar sempre o melhor aproveitamento possível das condições naturais desfavoráveis, ainda que apoiadas em base técnica frágil, utilizando na maior parte dos casos, tecnologias tradicionais.

**Figura 41** – Uso do Aradado manual



Fonte: Acervo do Projeto ELiHS.

**Figura 42** – Uso da Bruaca na colheita



Fonte: Acervo do Projeto ELiHS.

---

<sup>111</sup> As histórias narradas a respeito de sua origem são diversas, muito pouco se sabe sobre a sua formação e quem foram seus primeiros habitantes, dificultando o acesso a informações sobre seu princípio.

**Figura 43** – Batida de feijão manual



Fonte: Acervo do Projeto ELiHS.

As *Figuras* acima pertencem ao acervo pessoal do Projeto ELiHS e representam o trabalho agrícola desenvolvido, de forma ainda tradicional, na comunidade Volta do Angico, *locus* desta pesquisa. Vejamos mais algumas imagens representativas da Comunidade:

**Figura 44** – Retratos da Comunidade Volta do Angico



Fonte: Acervo do Projeto ELiHS.

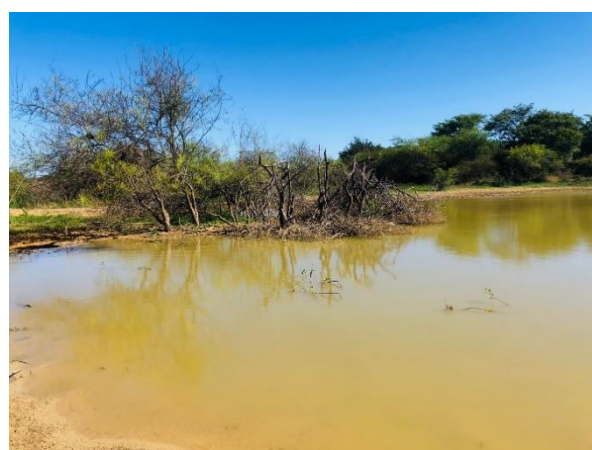
Essas poucas imagens, já tão representativas, marcam em seus detalhes a forma e a resistência de seu povo, mas afastados do processo de urbanização, mantêm um modo de vida simples e uma relação direta com o campo. Diante do contexto de invisibilidade das comunidades tradicionais, especialmente perante poderes públicos, a comunidade Volta do Angico ainda está condicionada a precárias condições no que tange, principalmente, a educação, saúde e cultura. Contudo, segundo relatos de moradores, muitos avanços já podem ser registrados hoje, uma vez que antigamente a comunidade não possuía recursos básicos, como água encanada e energia elétrica, a única forma de acesso à luz e à água era por meio do candeeiro (*cf.* Figura 45) e da cacimba (*cf.* Figura 46), cavada pelos próprios moradores para armazenamento de água. Atualmente já existe energia elétrica e uma caixa d'água que abastece a comunidade (*cf.* Figura 47).

**Figura 45** – Monitor do Projeto apresentando o Candeeiro



Fonte: Acervo do Projeto ELiHS.

**Figura 46** – Antiga cacimba – Volta do Angico



Fonte: Acervo do Projeto ELiHS.

**Figura 47** – Atual caixa de abastecimento de água



Fonte: Acervo do Projeto ELiHS.

**Figura 48** – Espaço para manutenção da água



Fonte: Acervo do Projeto ELiHS.

Apesar de ser uma Comunidade pequena e que usufrui de poucos recursos de órgãos públicos, sua população trabalha com o objetivo de resistir à opressão historicamente sofrida por seus ancestrais. Seus habitantes, a maioria produtores rurais, que vivem **no** e **do** campo, cultivam<sup>112</sup> a mamona (*cf.* Figura 49) e o milho (*cf.* Figura 50), sendo esses os principais percursos para a geração de renda na comunidade. Além disso, a influência da pecuária também se faz presente nas criações de ovinos, caprinos e suínos, bem como galinhas caipiras e galinhas de angola (cocá).

**Figura 49** – Terreiro de Mamona



Fonte: Acervo do Projeto ELiHS.

**Figura 50** – Sopros manual da Mamona



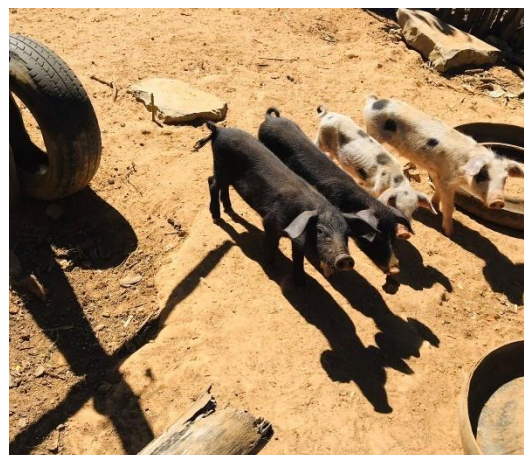
Fonte: Acervo do Projeto ELiHS.

**Figura 51** – Plantação de milho



Fonte: Acervo do Projeto ELiHS.

**Figura 52**– Criação de suínos



Fonte: Acervo do Projeto ELiHS.

<sup>112</sup> Há também a produção de abóbora, feijão e sorgo, no entanto em menor escala.

No que se refere à educação, segundo relatos de moradores, ainda é um ponto bastante precário e urgente, visto que o prédio destinado ao Ensino Fundamental I foi desativado no ano de 2008, sendo hoje apenas sede da Associação Comunitária de Moradores<sup>113</sup>. Tal fato induz que as crianças, jovens e adultos sejam direcionados a outra Comunidade vizinha, Mato Verde, ou à sede, Canarana, que fica localizada a 19 km, cenário que, no Brasil, remete a um processo de sabotagem de conquista de direitos, que perpassa por nossa própria história de formação enquanto sociedade.

**Figura 53** – Escola desativada – Volta do Angico



Fonte: Acervo do Projeto ELiHS.

Infelizmente a realidade da Comunidade Volta do Angico é replicada em tantas outras comunidades espalhadas pelo Território de Irecê – e porque não falarmos em Brasil. A escola pública, e de qualidade, não contempla parcela significativa da sociedade brasileira, em especial, as comunidades afro-brasileiras, fazendo com que o negro historicamente permaneça à margem do direito à educação. Neste sentido, não podemos deixar de relatar que no contexto das políticas afirmativas, temos a resolução nº 8, de 20 de novembro de 2012, que definiu as Diretrizes Curriculares Nacionais para a Educação Escolar Quilombola na Educação Básica, assegurando que:

Uma proposta de educação quilombola necessita fazer parte da construção de um currículo escolar aberto, flexível e de caráter interdisciplinar, elaborado de modo a articular o conhecimento escolar e os conhecimentos construídos pelas comunidades quilombolas. Isso significa que o próprio projeto político-pedagógico da instituição

<sup>113</sup>A comunidade possui associação desde 2015, através da qual tem sido possível a aquisição de políticas públicas que tem favorecido a permanência de seus membros em terras ainda pouco visibilizadas pelo poder público, em especial os setores de educação, saúde e cultura.

escolar ou das organizações educacionais deve considerar as especificidades históricas, culturais, sociais, políticas, econômicas e identitárias das comunidades quilombolas, o que implica numa gestão democrática da escola que envolve a participação das comunidades escolares, sociais e quilombolas e suas lideranças. Por sua vez, a permanência deve ser garantida por meio da alimentação escolar e a inserção da realidade quilombola em todo o material didático e de apoio pedagógico produzido em articulação com a comunidade, sistemas de ensino e instituições de Educação Superior (Brasil, 2012, p. 26).

Fechar escolas é retroceder, é sucumbir direitos adquiridos. Portanto, é preciso lutar não apenas pelo direito ao acesso à educação, direito humano básico para formação da pessoa, mas principalmente por condições de permanência.

No que tange à cultura, é possível relatar a permanência de algumas festas tradicionais, a saber: a festa de reis ou reisado (comemorado em dezembro); a festa dos Caretas (comemorado em abril); Carurus; e eventos esportivos, como campeonatos de futebol e cavalgadas.

**Figura 54** – Manifestação popular: Festa dos Caretas abril/2023



Fonte: Acervo do Projeto ELiHS.

Os retratos apresentados até o momento refletem um modo de vida simples, de disparidades socioeconômicas/sociais, que não representam apenas um simples legado do passado, haja vista ainda estar sendo perpetuado pela estrutura desigual de oportunidades sociais a que as comunidades tradicionais, em especial as afro-brasileiras, estão expostas no presente.

Atualmente residem na comunidade em média 277 membros, dos quais 90% se consideram negros, em sua maioria adultos, com idades que variam entre 31 e 59 anos. No que tange a questão religiosa, predomina entre os moradores locais o catolicismo, no entanto, há 01 (uma) igreja protestante (cf. Figura 55) e 01 (um) terreiro de umbanda<sup>114</sup>.

**Figura 55** – Igreja protestante – Volta do Angico



Fonte: Acervo do Projeto ELiHS.

Ademais, a comunidade conta com acesso precário ao telefone, entretanto, cerca de 40% dos moradores possuem conexão à internet. Como já retratado, em relação à educação, apenas crianças do Fundamental I são assistidas pela comunidade vizinha, Mato Verde, e menos de 50% dos moradores concluíram o ensino médio<sup>115</sup>.

Este breve retrato do povoamento de Volta do Angico traz consigo a tentativa de ratificar que a mesma se caracteriza como uma comunidade tradicional afro-brasileira [-urbanizada], [+ afrodescendente] e [- informatizada], caracterizando dentro dos padrões metodológicos propostos para seleção da comunidade, *locus* da pesquisa de análise de dados linguísticos. Sendo assim, acreditamos que os resultados a serem encontrados se situem nas frequências de uso que podem revelar uma interferência mais direta do contato entre línguas ou “[...] *vis-à-vis* marcas diluídas que refletem residualmente uma influência indireta do contato

<sup>114</sup> O Projeto não possui fotos em seu acervo.

<sup>115</sup> Informações adquiridas através do questionário respondido pelo Presidente da Associação dos Moradores.

entre línguas” (Lucchesi, 2000, p.57), como já demonstrando por Andrade (2003) para a comunidade de Helvécia, na Bahia.

### 5.2.2 Procedimentos para a análise estatística

As discussões suscitadas até o momento nos permitem afirmar que os processos de variação/mudança estão inter-relacionados a fatores tanto de natureza linguística, quanto extralinguística (Weinreich; Labov; Herzog, 2006 [1968], p. 126). À luz dos preceitos labovianos é, portanto, necessário considerar fatores atuantes sobre fenômenos linguísticos variáveis, na tentativa de verificarmos os efeitos de variáveis independentes, posto que “[...] as pressões internas, estruturais, e as pressões sociolinguísticas agem em alternância sistemática no mecanismo da mudança linguística” (Labov, 2008 [1972], p. 214).

Portanto, a comunidade de Volta do Angico, situada no município de Canarana, representará a amostra do português popular de comunidade afro-brasileira nas análises a serem descritas e comentadas em seção posterior, diante dos critérios socioculturais descritos no *Quadro 22*.

**Quadro 22** – Descrição da população amostra do PPB representativo da Comunidade de Volta do Angico

Inf.	Sigla de identificação	Sexo	Faixa Etária*	Profissão	Tempo de gravação
1	F. S. N.	M	I	Agricultor	54min. 38seg
2	V. O.	M	I	Agricultor	1h 10min e 46seg
3	J. A. A. S.	M	II	Agricultor	43 min. 26 seg.
4	J. S. B	M	II	Agricultor	1h e 06 seg.
5	J. Q.	M	III	Aposentado	50min. 19seg.
6	D. F. S.	M	III	Aposentado/ Agricultor	1h. 02min. 46seg.
7	T. J. S.	F	I	Agricultora	38min. 5seg.
8	J. P. S.	F	I	Agricultora	1h. 01min. 18seg
9	M. L.	F	II	Aposentada	40min. 13seg.
10	J.A.S.	F	II	Agricultora	1h. 02min. 28seg
11	M. R. O	F	III	Aposentada	1 hora 16min. 18seg.
12	F. P. B.	F	III	Aposentada	40min. 02seg.

\*As faixas etárias I, II e III correspondem, respectivamente, aos informantes até 39 anos, 40 a 59 anos, a partir de 60 anos.

Fonte: Elaborado pela autora.



Lopes (2001, p. 114) acredita que em uma pesquisa sociolinguística, a seleção dos participantes<sup>116</sup> deve refletir as hipóteses iniciais e orientar a constituição de uma amostra representativa das variáveis sociais em estudo. Em vista disso, no que tange aos 12 (doze) participantes que compõem a amostra, todos se encontram na faixa etária de 18 a 82 anos e possuem, no máximo, fundamental incompleto (até o 5º ano), caracterizando a amostra enquanto representativa de um português popular. Quanto à origem dos falantes, todos são nascidos e criados em Volta do Angico e nenhum deles saíram da comunidade por mais de 6 meses consecutivos.

A amostra a ser analisada perfaz 9 horas, 58 minutos e 25 segundos de gravação, que foram transcritas e disponibilizadas eletronicamente através do *site* do Projeto ELiHS, compondo o banco de dados com uma amostra representativa do Português Popular de comunidade afro-brasileira do Território de Irecê. A descrição metodológica, da gravação à transcrição dos áudios, já foi detalhadamente descrita na caracterização metodológica do Projeto ELiHS (*cf.* seção 4.2.2), não sendo necessário, neste momento, retomarmos.

Ainda em relação a questões metodológicas é importante declarar que o fenômeno da concordância de número no sintagma nominal pode ser analisado sob duas perspectivas, atomística e não atomística. Na perspectiva atomística considera-se como dado de análise cada um dos constituintes flexionáveis dos SNs plurais. Destarte, se o constituinte estiver formalmente marcado, considera-se presença de concordância, caso contrário, considera-se ausência. Porém, Scherre (1988, p. 64) chama atenção para o termo ‘concordância’, uma vez que não o considera apropriado a todas as situações, já que muitas vezes apenas um elemento do sintagma é formalmente marcado, como em (9), outras vezes tem-se SNs sem nenhuma marca formal de plural, como em (10). Por isso, a autora prefere tratar o fenômeno como indicação de pluralidade e não concordância, já que a “concordância gramatical implica harmonia formal em pelo menos dois elementos de uma dada construção” (SCHERRE, 1988, p. 64), o que não ocorre em (9) e (10).

(9) **As menina** foi lá buscar e não pegou **as brusa**. (V. O.)

(10) Quando era **cinco hora**, a gente já ta no pouso de novo. (D. F. S.)

---

<sup>116</sup> Assumimos a terminologia participantes diante das prerrogativas do Comitê de Ética em Pesquisa e por entender que esses não apenas “informam” dados, mas participam desta pesquisa através dos seus relatos e composição para acervo oral do Projeto, um patrimônio imaterial.

Nesta perspectiva de análise as diversas propriedades linguísticas relativas a cada um dos elementos nominais da amostra são analisadas cuidadosamente e o linguista pode trabalhar com um conjunto expressivo de variáveis<sup>117</sup>, buscando as possíveis adequações ao seu objeto de estudo ou fenômeno em análise.

Já na perspectiva não atomística – conhecida também como sintagmática –, assumida nesta tese, analisa-se o SN como um todo, não havendo análises individuais dos elementos flexionáveis do SN, como propõe a perspectiva atomística. Sendo assim, leva-se em conta “[...] a presença de marca em todos os elementos flexionáveis do SN *versus* ausência de marca em pelo menos um dos elementos flexionáveis do SN” (Scherre, 1994, p. 39), considerando como aplicação da regra, a presença de plural formal em todos os elementos flexionáveis que constituem o SN, como em (11), e como não aplicação, a falta de pelo menos uma marca formal de plural, como em (12).

(11) Se eu pudesse eu tava cuidando **das prantas**.

(12) Ele é que sabe **as históriaø**.

Sendo assim, define-se, nesta pesquisa, como variável dependente<sup>118</sup> a presença ou ausência de marca de plural nos elementos do sintagma nominal. Portanto, observamos nas amostras dois padrões básicos e opostos:

- (i) Uma variante redundante, em que se tem a marca de plural (morfema -s) em todos os constituintes flexionáveis do SN;
- (ii) Outra, simplificada, em que não se tem a marca de plural em todos constituintes do SN.

A fim de observar de que forma a variação no uso da regra de concordância se encaixa na estrutura linguística e social do dialeto da comunidade afro-brasileira Volta do Angico, tomamos como hipótese inicial que a presença de marca pode estar relacionada a 10 (dez) variáveis, 02 (duas) socioculturais e 06 (seis) linguísticas ou estruturais, as mesmas utilizadas

---

<sup>117</sup> No trabalho, na perspectiva atomística, sob a variação na concordância de número no SN, costuma-se utilizar as seguintes variáveis, sendo estas utilizadas por Scherre (1988), a citar: processos morfofonológicos de formação do plural; tonicidade dos itens lexicais singulares; número de sílabas dos itens lexicais singulares; posição linear do elemento no SN; classe gramatical do elemento nominal; marcas precedentes ao elemento nominal analisado; contexto fonético/fonológico seguinte ao elemento nominal sob análise; função sintática do SN (codificada em cada um de seus constituintes); animacidade dos substantivos; grau dos substantivos e dos adjetivos e formalidade dos substantivos e dos adjetivos.

<sup>118</sup> Chama-se variável dependente o fenômeno que está sendo estudado. Considera-se dependente por ter como pressuposto que a presença de marca não é aleatória, mas condicionada por (ou dependente de) outras variáveis, as independentes (Lopes, 2001, p. 129).

por Andrade (2003) na Comunidade de Helvécia, inserindo em nossas análises 02 (duas) variáveis aleatórias<sup>119</sup>.

**Quadro 23** – Descrição das variáveis resposta e predictoras

<b>VARIÁVEL RESPOSTA</b>	Concordância Nominal de Número	
<b>VARIÁVEIS PREDITORAS</b>	<b>Linguísticas</b>	Configuração sintagmática do SN
		Função sintática do SN
		Posição do SN em relação ao verbo
		Número de constituintes absolutos do SN
		Número de constituintes flexionáveis do SN
		Saliência fônica
	<b>Socioculturais</b>	Sexo de nascimento
		Faixa etária
<b>VARIÁVEIS ALEATÓRIAS</b>	Item lexical	
	Participante	

Fonte: Elaborada pela autora.

Após levantamento de 1253 ocorrências, observamos, empiricamente, que a variável não marcada se apresentava quase que categoricamente, o que implicaria diretamente na utilização do pacote estatístico *Goldvarb X*<sup>120</sup>, proposto no projeto e na qualificação desta tese, que realiza uma análise multivariada amplamente empregada em estudos de variação linguística, hoje ainda o mais utilizado nas análises quantitativas de cunho Sociolinguístico.

Desde a criação do *Variable Rule Analysis*<sup>121</sup> (VARBRUL), muitas foram as alterações realizadas no programa para que esse se tornasse cada dia mais prático, confiável e com mais ferramentas operacionais, o que tem colaborado na aplicação de testes, cruzamentos de dados, no entanto, ainda com muitas limitações. Assim, nós, linguistas, temos nos deparado com contextos variáveis que têm exigido mais opções para análises e, é nesse contexto que se

<sup>119</sup> As variáveis aleatórias se classificam como níveis não controlados pelo pesquisador.

<sup>120</sup> O *Goldvarb X* é um software gratuito, desenvolvido por Sankoff, Tagliamonte e Smith (2005), para análise de *corpus*, executável em versões para o *Windows*, *Linux* e *Macintosh*. Seu arquivo tem aproximadamente 2,66MB, por isso é considerado leve e dispensa a necessidade de instalação, o que torna a utilização mais fácil mesmo para usuários iniciantes. O programa está disponível na página <http://individual.utoronto.ca/tagliamonte/goldvarb.html>.

<sup>121</sup> Análises de regras variáveis.

encontra a linguagem de programação denominada *R* (R CORE TEAM, 2020)<sup>122</sup>, amplamente conhecido, gratuito e aberto a livre acesso, com seus respectivos *scripts*, em que “[...]o usuário não está restrito a opções pré-programadas por um *software* (como o GoldVarb X, SPSS, Excel etc.), *cujos programadores dificilmente podem ter previsto tudo o que um usuário pode ter interesse em executar*” (Oushiro, 2022, p. 16, grifo nosso). Nesse sentido, a autora ainda afirma que

O GoldVarb X é uma ferramenta útil, contanto que se saiba o que ela é capaz de fazer e o que não é capaz de fazer. *É um programa feito para realizar apenas um tipo de análise: modelos multivariados fixos de regressão logística com variáveis nominais* – ou seja, a variável dependente deve, necessariamente, ser binária (variante A vs. variante B, como “concordância nominal padrão” vs. “concordância nominal não padrão”),<sup>7</sup> e as variáveis independentes não podem ser contínuas (como, p.ex., idade do falante: 20, 22, 27 anos), mas devem ser variáveis nominais (como, p.ex., faixas etárias: de 20 a 35, de 36 a 50 etc.). No GoldVarb X, não é possível modelar, adequadamente, variáveis contínuas [...]; variáveis dependentes/resposta enérgicas, com mais de dois níveis de variantes; *o papel de efeitos aleatórios, próprios de cada amostra, como falantes individuais*; não se pode visualizar os dados em diferentes tipos de gráficos; e não se pode comparar as medidas geradas com resultados de outros estudos, em outros programas, pois os pesos relativos são uma medida usada apenas por sociolinguistas (Oushiro, 2022, p. 17-18, grifo nosso).

Os fatores já levantamos por Oushiro (2022) nos levaram a conhecer a linguagem de programação em cena e decidimos sobre a viabilidade de efetuarmos as análises estatísticas através da regressão logística, por meio de modelos lineares generalizados de efeitos fixos e mistos, obtidos com o *R*<sup>123</sup> (R CORE TEAM, 2020), por meio de uma de suas *interfaces* o *RStudio*<sup>124</sup>, usando-se a função *glmer* do pacote *lme4*, cujos resultados obtidos são exibidos e comentados na próxima seção.

Este método estatístico produz, a partir de um conjunto de observações, um modelo que licencia a predição de valores tomados por uma variável categórica, comumente binária, através de um encadeamento de variáveis preditoras contínuas e/ou binárias. Este tipo de abordagem quantitativa de dados pode desvelar padrões de variação que escapariam à vista, principalmente quando se trabalha com milhares de ocorrências codificadas de acordo com parâmetros socioculturais e linguísticos, como a que nos propomos nesta tese.

<sup>122</sup> Como afirma a professora Livia Oushiro (2022, p. 9), infelizmente, até o momento, a formação em Estatística não é componente obrigatório da maior parte dos cursos de Letras e de Linguística, e tem de ser buscada por iniciativa própria de pesquisadores, de pós-graduação ou graduação, principalmente no que tange a Plataforma R. Por isso, contamos o auxílio do Prof. Dr. Elias de Souza Santos, que orienta e publica resultados de estudos cujas modelagens são obtidas por meio da linguagem de programação denominada R (R CORE TEAM, 2020).

<sup>123</sup> Disponível em: <https://www.r-project.org/>.

<sup>124</sup> O *Rstudio*, que utilizaremos em nossa análise estatística, é um dos vários ambientes de desenvolvimento integrado do *R*, que apresenta uma interface mais funcional e opera através da execução de *scripts* computacionais, exibindo funções para a distribuição de probabilidade, testes estatísticos, produção de gráficos (GRIES, 2013). Disponível em: <https://www.rstudio.com/products/rstudio/>

Após termos tomado a decisão em relação ao uso da linguagem de programação em perspectiva, submetemos as ocorrências a uma tabulação em planilha de Excel, salva em formato .csv, para ser carregada no R. As planilhas foram organizadas segundo a **variável-resposta** – registrando a aplicação, ou não, da marcação de plural em todos os itens pluralizáveis do SN –, e as **variáveis previsoras** – a cada dado a ser considerado na análise estatística –, seguindo a orientação de Oushiro (2022, p. 14, grifo da autora) que recomenda seguirmos “[...] o formato ‘caso por variável’ apresentado por Gries (2019, p. 29–30) – também chamado de formato *longtable*”, conforme *Figura 59*:

**Figura 59** – Tela de codificação no formato “caso por variável” (Oushiro, 2022)

	A	B	C	D	E	F	G	H	I	J	K
1	Con.nominal	Sexo	Fx.etária	Conf.sint.SN	Pos.SN.verbo	Núm.const.SN	Núm.const.flex.SN	Sal.fônica	Item.léxico	Participante	Ocorrência
2	Ausência	Feminino	Um	QT+N	Direita	Dois	Dois	Regular	Coisa	T.J.S	não tinha MUITAS COISA
3	Ausência	Feminino	Um	QT+N	Direita	Dois	Dois	Regular	Coisa	T.J.S	hoje já tem VÁRIAS COISA
4	Ausência	Feminino	Um	QT+N	Posição indistinta	Dois	Dois	Regular	Coisa	T.J.S	MUITAS COISA
5	Ausência	Feminino	Um	QT+N	Posição indistinta	Dois	Dois	Regular	Coisa	T.J.S	MUITAS COISA
6	Ausência	Feminino	Um	DET+N	Esquerda imediata	Dois	Dois	Regular	Pessoa	T.J.S	AS PESSOA é camarada
7	Ausência	Feminino	Um	QT+N	Direita	Dois	Dois	Regular	Coisa	T.J.S	brincava de chicot... VÁRIAS COISA
8	Ausência	Feminino	Um	DET+N	Direita	Dois	Dois	Regular	Pedrinha	T.J.S	brinca com UMAS PEDINHA
9	Presença	Feminino	Um	DET+N	Esquerda imediata	Dois	Dois	Regular	Outros	T.J.S	ficava um em pé no poste e OS OUTROS ia escondê
10	Presença	Feminino	Um	NUM+N	Posição indistinta	Dois	Um	Regular	Horas	T.J.S	De sete até OITO HORAS
11	Ausência	Feminino	Um	P+DET+N	Direita	Dois	Dois	Regular	Aula	T.J.S	Eu gostava DAS AULA dele
12	Ausência	Feminino	Um	DET+N	Posição indistinta	Dois	Dois	Regular	Coisona	T.J.S	É né AQUELAS COISONA não, mas vai...té que presta
13	Ausência	Feminino	Um	DET+QT+N+A	Posição indistinta	Três	Três	Regular	Tempero	T.J.S	OS OTOS TEMPERO NECESSÁRIO
14	Ausência	Feminino	Um	NUM+N	Posição indistinta	Dois	Um	Irregular	Mês	T.J.S	Um de um ano e CINCO MÊS
15	Presença	Feminino	Um	NUM+N	Posição indistinta	Dois	Um	Regular	Cinco	T.J.S	um de um ano e cinco mês e oto de CINCO ANOS.
16	Ausência	Feminino	Um	QT+N	Direita	Dois	Dois	Regular	Coisa	T.J.S	tinha MUITAS COISA
17	Ausência	Feminino	Um	DET+QT+N	Direita.imediata	Três	Três	Regular	Velho	T.J.S	Tem muito menino respondão...responde OS MAIS VELH
18	Ausência	Feminino	Um	DET+N	Direita.imediata	Dois	Dois	Regular	Ropa	T.J.S	Só vendo UMAS ROPA
19	Ausência	Feminino	Um	DET+NUM+N	Direita.imediata	Três	Três	Regular	Ano	T.J.S	tem UNS TRÊS ANO
20	Ausência	Feminino	Um	DET+N	Posição indistinta	Dois	Dois	Irregular	Condições	T.J.S	UMA CONDIÇÕES melhor né

Fonte: Elaborada pela autora.

Como observado na *Figura 59*, nessa estratégia os dados são tabulados da seguinte forma: (i) a primeira linha contém os nomes das variáveis; (ii) cada linha representa um e apenas um dado (uma observação da variável dependente); (iii) cada uma das colunas representa uma e apenas uma variável.

Após o processo de tabulação, efetuamos análises estatísticas de regressão logística, por meio de modelos lineares generalizados de efeitos fixos e mistos. Iniciamos com a análise univariada na tentativa de verificarmos o comportamento das variáveis predictoras, por meio do teste de *qui-quadrado*, o qual tende a indicar as variáveis que apresentam ou não significância estatística, podendo essa última realidade, por decisão do analista, ser consequentemente excluída dos modelos de efeitos fixos e mistos.

Com o teste de *qui-quadrado*, excluimos as variáveis *Configuração sintagmática do SN*, *Função sintática do SN* e *Número de constituintes absolutos do SN*, por elas não terem

apresentado significância estatística, redimensionando, assim, nosso quadro de variáveis preditoras (cf. Quadro 24).

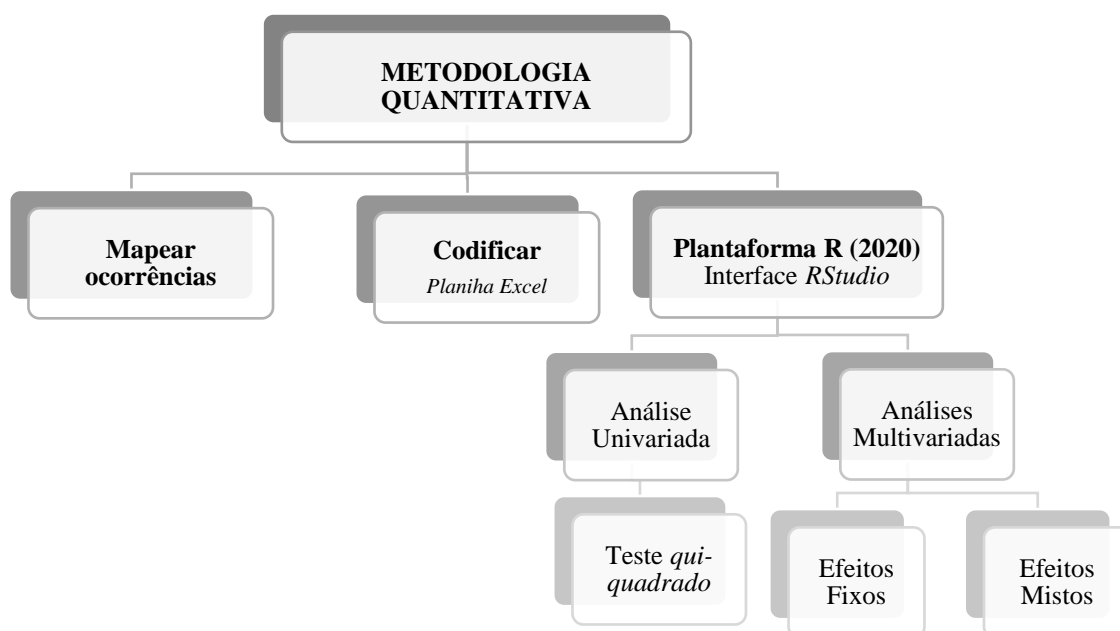
**Quadro 24** – Descrição das variáveis preditoras linguísticas, após teste *qui-quadrado*

Variáveis Preditoras Linguísticas
Função sintática do SN
Posição do SN em relação ao verbo
Número de constituintes flexionáveis do SN
Saliência fônica

Fonte: Elaborada pela autora.

Ajustados os dados, após teste de *qui-quadrado*, iniciamos análises de efeitos fixos, objetivando identificar as ocorrências variáveis da concordância nominal de número, sem o controle das variáveis aleatórias – participante e item lexical. Os resultados obtidos na análise de efeitos fixos subsidiaram a análise contrastiva com os resultados da análise de efeitos mistos, que foram realizadas na sequência, para certificação de que a estatística dos modelos se mantivessem constantes em ambos – fixos e mistos. Por conseguinte, esses são os passos metodológicos aplicados à análise estatística, resumido na Figura 60 a seguir.

**Figura 57** – Passos metodológico no uso da Plataforma R



Fonte: Elaborada pela autora

Aqui buscamos apresentar, ainda que sinopticamente, os procedimentos metodológicos adotados durante a fase de levantamento e análise dos dados desta pesquisa, compreendendo que

a utilização de recursos estatísticos em análises sociolinguísticas não é mero recurso metodológico; ela na verdade deriva do postulado teórico que Weinreich, Labov e Herzog (2006 [1968]) estabeleceram para o exame da mudança linguística: não há variação livre. Isso significa que toda variação presente no sistema linguístico será condicionada por fatores linguísticos ou sociais que deverão ser identificados pelo linguista. Em muitos casos, como mostrou Labov (2008 [1972]), esses fatores condicionantes não são facilmente identificáveis, diante da intrincada relação entre diversas forças que atuam no sistema. Como essas forças não se devem ao acaso, elas precisam ser identificadas com precisão – aí é que entram as análises estatísticas, já que elas permitem estabelecer os padrões de variação de cada comunidade de fala. (Santos, 2012, p. 145).

Assim, no encontro entre teoria e métodos fomos direcionados a novas escolhas, a dos recursos computacionais que favorecessem a concretização efetiva dos objetivos fulcrais da pesquisa. Neste campo de escolhas e enquadramentos metodológicos optamos pelo *Rstudio*, entendendo que essa linguagem de programação atende ao cumprimento de todas as etapas da pesquisa de cunho sociolinguístico, e para a análise empírica que propomos realizar, atende as variáveis socioculturais e linguísticas elencadas como possíveis condicionantes do processo de variação do fenômeno em análise. Assim sendo, as análises que seguem embasam-se nos pressupostos da Teoria da Sociolinguística Quantitativa (Labov, 2008 [1972]), com o emprego do *R* (R CORE TEAM 2020) para seus respectivos processamentos estatísticos.

### **5.3 Análise estatística da concordância nominal de número em amostra da comunidade afro-brasileira Volta do Angico**

Seguindo a metodologia descrita em seção anterior, aqui discutiremos os dados estatísticos obtidos através da linguagem de programação denominada *R* (R CORE TEAM, 2020), por meio de sua interface o *Rstudio* (EQUIPE RSTUDIO, 2020), com os pressupostos da Teoria da Variação e Mudança (Weinreich; Labov; Herzog, 2006) e da Transmissão Linguística Irregular (Lucchesi, 2009, 2008, 2003), com os quais acreditamos ser possível compreender o processo de variação na concordância nominal na variedade do português falado na comunidade afro-brasileira Volta do Angico. Dessa forma, pretendemos apontar os fatores que podem motivar as marcas explícitas de plural nos sintagmas nominais e, em última instância, para os debates sobre as origens do Português do Brasil (Lucchesi, 2009, 2008, 2003, 1998; Naro, 2003; Naro e Scherre, 2007).

Assim seguiremos apresentando os resultados obtidos em relação a variável resposta e as variáveis preditoras inclusas na análise da marcação de plural nos 1.253 sintagmas nominais, presentes na amostra de análise.

### 5.3.1 Análises multivariadas

Como já relatado, por variável resposta entende-se o fenômeno que se pretende investigar, a variação na concordância nominal de número. À face disso, fica estabelecido como variável de resposta, que a aplicação da regra ocorre se houver presença de marcas em todos os itens pluralizáveis do SN, como em (13), enquanto que a ausência dessa marca, seja em apenas um dos elementos do SN sob análise, constitui-se falta de concordância, como em (14).

(iv) Inserção da marca de plural em todos os elementos do SN

(13) “Assim **os alunos** assim às vez quando chove eles perde aula.” (J. A. S.)

(v) Não marcação de alguns elementos do SN

(14) “Cada qual tem **os seus emprego**Ø<sup>125</sup>.” (J. Q.)

Na *Tabela 5* podemos visualizar os resultados obtidos, através da análise quantitativa dos 1253 SNs mapeados na amostra.

**Tabela 5** – Taxas de uso da concordância no SN em amostra da comunidade Volta do Angico

VARIÁVEL RESPOSTA	Nº OCORRÊNCIAS	%
<b>SN com todas as marcas de plural</b>	131	10,45
<b>SNs sem alguma das marcas de plural</b>	1122	89,55

Fonte: Elaborada pela autora, com dados da pesquisa.

Os resultados indicam que na amostra há uma tendência significativa ao uso não canônico na indicação do plural nos SNs (89,54%) em relação ao uso canônico, com as marcas

<sup>125</sup> A simbologia Ø utilizada neste trabalho, reflete a confluência formal entre o zero plural e a morfologia do singular, caracterizada pelo morfema zero, como define Scherre e Naro (2006, p. 107).



de plural em todos os itens do SN (10,45%). Este resultado nos permite afirmar, ainda em uma análise preliminar, que embora a tendência do português brasileiro seja a marcação explícita de número nos elementos que compõem o SN (Braga, 1977; Lemle; Naro, 1977; Naro; Scherre, 1991, 1993, 2007; Scherre, 1978, 1988, 1994, 1998; Scherre; Naro, 1992, 1993, 1998, 2006; entre muitos outros), ainda há comunidades linguísticas, como a comunidade afro-brasileira Volta do Angico, Bahia, que não conservam a atual variante prototípica, de marcas explícitas. A razão disso pode estar relacionada à fatores socioculturais, formação sócio-histórica da comunidade auto identificada como remanescente de quilombos, junto a isso às particularidades do município e condição de vulnerabilidade social a qual a comunidade se encontra, cabendo o destaque para o fato de menos de 50% da comunidade não terem concluído o Ensino Médio, concomitante a isso, não terem mais sede escolar na Comunidade, sendo obrigados a se deslocarem diariamente 16,3 km, até a sede, ou à comunidade vizinha Mato verde.

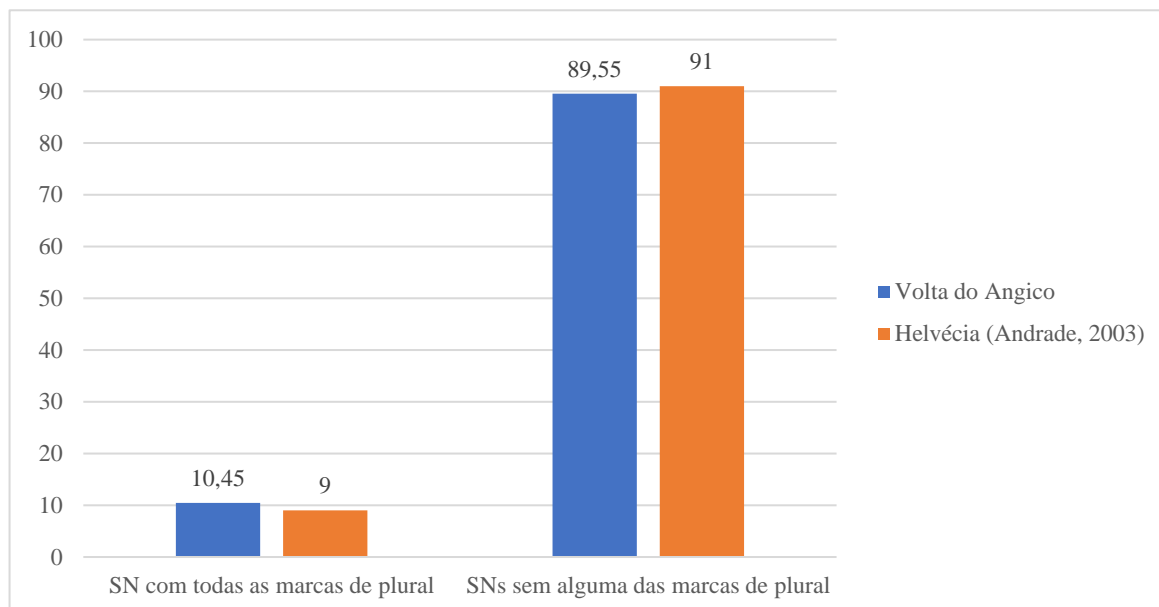
Lembramos que os resultados são representativos, composto por um *corpus* de 12 participantes, os quais foram selecionados de modo aleatório, sendo estratificados conforme critérios já estabelecidos na seção metodológica. Neste contexto, ainda cabe frisar que em relação ao fenômeno investigado (a variação na marcação de número no SN), pode-se presumir duas forças atuantes nos resultados apontados na *Tabela 5*, a saber:

- (i) uma em direção à **ruralidade** em função das raízes afro-brasileiras da comunidade, visto que a comunidade está situada à 16,3 km de sua sede Canarana e 61,9 km da cidade de Irecê (centro comercial), conservando atividades econômicas do setor primário, em especial a agricultura (*cf.* seção 5.2.1). Ainda cabe suscitar a ideia de que em decorrência de os falantes manterem o seu círculo de interação restrito à comunidade – com ligação direta com o meio rural –, se sentirem pertencentes a sua comunidade/território e/ou não almejem mudança para um meio mais urbano, direciona-os à um contexto ainda mais rural;
- (ii) outra em direção à **escolaridade**, visto que todos os participantes não são escolarizados (analfabetos ou semiescolarizados), configurando o *corpus* enquanto representativo do português popular.

Os resultados estatísticos apresentados e seus possíveis argumentos explicativos estão, sem dúvidas, condicionados à existência de uma realidade sociolinguística bipolarizada, retrato de um panorama de polarização sociolinguística presente no Português Brasileiro (Lucchesi, 2001, 2002, 2006) e ratificada nos dados da comunidade Volta do Angico. Mesmos cientes de

uma realidade presente, em especial, no Português Popular, faz importante averiguar se há conformidade entre os resultados obtidos por Andrade (2003), que também pesquisou o fenômeno da variação na concordância nominal de número na comunidade Helvécia, caracterizada enquanto comunidade rural afro-brasileira.

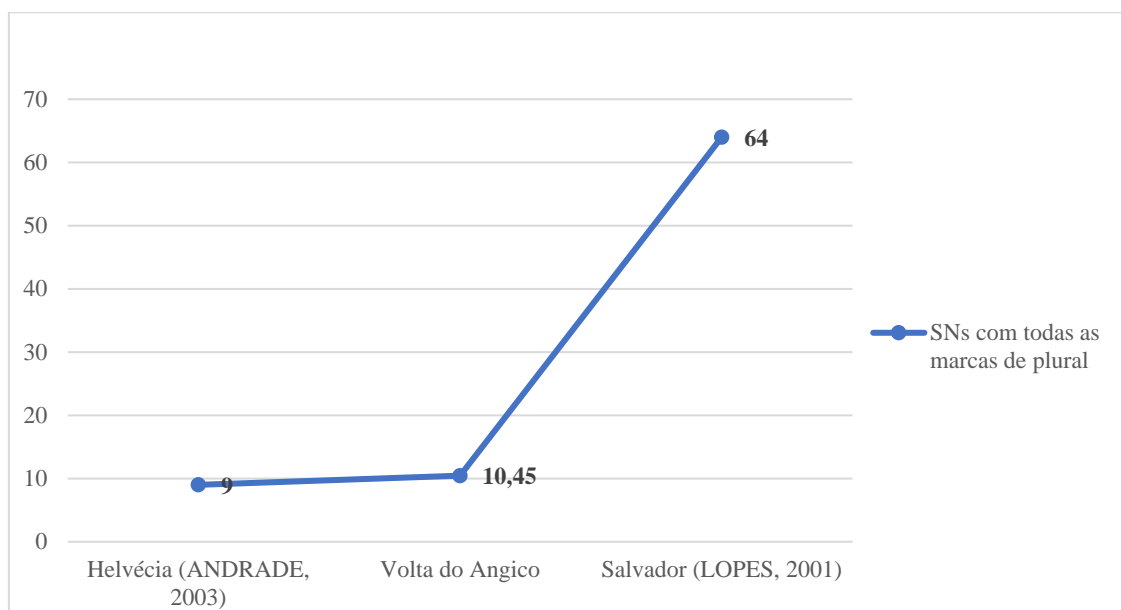
**Gráfico 14** – Resultado contrastivo entre as comunidades Volta do Angico e Helvécia (Andrade, 2003)



Fonte: Elaborado pela autora.

Considerando os resultados representados no *Gráfico 14*, algumas questões já podem ser destacadas, a saber: mesmo a comunidade Volta do Angico não se caracterizando em termos de [+ isolada], assim como Helvécia, temos um resultado estatístico aproximativo no que tange a não marcação de plural no SN (89,55% para Volta do Angico e 91% para Helvécia), reafirmando o posicionamento de Lucchesi (2009) sobre a ideia da perda de morfologia flexional no PB popular ser um indicador da influência do contato linguístico. Para seguirmos com a análise contrastiva inseriremos os dados de Lopes (2001), que analisa a variação da concordância de número no SN nas falas popular e culta na cidade de Salvador, capital baiana, no entanto, aqui centraremos nossos olhares aos dados do português popular, dos participantes que cursaram até o 5º ano de escolarização, perfil escolar semelhante aos da comunidade de Volta do Angico e Helvécia (Andrade, 2003).

**Gráfico 15** – Resultado contrastivo entre as comunidades Volta do Angico, Helvécia (ANDRADE, 2003) e Salvador (LOPES, 2001)



Fonte: Elaborado pela autora.

A leitura do *Gráfico 15* nos permite inferir que mesmo ao considerarmos as comunidades afro-brasileiras como rurais e com rede social isolada<sup>126</sup> (Bortoni-Ricardo, 2021) na contramão da cidade de Salvador que se caracteriza como urbana e com rede social integrada<sup>127</sup> (Bortoni-Ricardo, 2021) temos resultados que demonstram que a variação no processo de concordância faz parte do vernáculo de quase a totalidade das populações, independentemente de serem caracterizadas como rurais, rurbanas ou urbanas, o que caracteriza a variação na concordância nominal de número como processo inerente à gramática da língua portuguesa no Brasil. O ponto que merece destaque centra-se no quantitativo estatístico, visto que as comunidades rurais, Volta do Angico (10,5%) e Helvécia (9%), apresentam um proporção de marcação de plural nos SNs que representa um processo de perda morfológica, caminhando para um sistema sem concordância. Como afirma, Lucchesi (2008, p. 373, grifo do autor),

dentre os níveis mais afetados na estrutura gramatical do português brasileiro pela transmissão linguística irregular deve-se destacar a significativa erosão provocada na morfologia flexional do nome e do verbo e a fortiori nas regras de concordância que se apoiam nesses morfemas [...]. O amplo quadro de variação na concordância nominal e verbal decorrente desse processo de transmissão linguística irregular de tipo leve constitui uma verdadeira linha demarcatória na realidade linguística brasileira,

<sup>126</sup> Rede de contatos estrito à família extensas e vizinhos.

<sup>127</sup> Rede de contato com pessoas em muitos contextos sociais.

dividindo-a em duas grandes normas: a *norma culta*, só indiretamente afetada por esse processo, e a *norma popular*, cuja formação é indelevelmente marcada por processos de mudança induzidos pelo contato entre línguas [...].

Essas discussões iniciais, e estudos já realizados (*cf.* seção 5.1.1), demonstram que a regra de concordância de número no interior do SN oferece um quadro bastante significativo no PB, portanto, seguiremos com as análises das variáveis preditoras socioculturais e linguísticas, as quais demonstrarão diferentes medidas que, porventura, poderão indicar novos caminhos na compreensão dos resultados obtidos e em relação ao atual estágio do uso da concordância nominal de número, na variedade português falado em Volta do Angico, em comparação com Helvécia (Andrade, 2003).

#### 5.3.1.1 Análise multivariada de efeitos fixos e mistos

Os dados apresentados até o momento indicam, preliminarmente, uma direção ao processo de inovação, tornando a não marcação de plural no SN uma marca prototípica na comunidade de fala Volta do Angico, muito semelhante ao que foi identificado na comunidade de Helvécia (Andrade, 2003). Todavia, esses dados gerais precisam ser analisados estatisticamente para compreendermos as forças que condicionam essa alteração prototípica, desse modo, seguiremos analisando e descrevendo as possíveis influências de forças das variáveis previsoras, linguísticas e socioculturais, diante da análise de regressão logística de efeito fixos e mistos.

Compreendendo que “a decisão sobre quais variáveis têm efeito de fato para descrever, explicar e prever o comportamento da variável resposta” (Oushiro, 2021, p. 290), cada variável previsora foi tomada diante de uma hipótese a respeito da variável resposta, e as hipóteses, portanto, têm por fundamento a teoria e a literatura já publicada sobre o fenômeno, tomando como base comparativa, nesta tese, a pesquisa de Andrade (2003), sobre a variação na concordância nominal de número em *corpus* oral da comunidade afro-brasileira Helvécia. Vejamos nossos resultados, iniciando com um modelo multivariado de regressão logística de efeitos fixos.

**Tabela 6** – Estimativas dos parâmetros do modelo (de regressão logística, modelo linear generalizado de efeitos fixos com interação) da concordância nominal de número

	Apl./N	Estimativa	Erro Padrão	Valor Z	(p)
<b>Sexo</b>					
Masculino (V. ref.)	83/586 (7%)				
Feminino	48/667 (4%)	-0.1088	0.2705	-0.402	0.687534
<b>Faixa Etária</b>					
Um (V. ref.)	81/357 (7%)				
Dois	33/429 (3%)	-0.5999	0.2643	-2.270	0.023225 *
Três	17/467 (1%)	-3.0134	0.6084	-4.953	7.31e-07 ***
<b>Nº de Constituintes Flexionáveis do SN</b>					
Dois (V. ref.)	95/926 (8%)				
Três	05/95 (01%)	-1.0358	0.4820	-2.149	0.031657 *
Um	31/219 (2%)	0.7063	0.2437	2.898	0.003754 **
<b>Saliência Fônica</b>					
Irregular (V. ref.)	28/184 (02%)				
Regular	103/1069 (8%)	-0.4985	0.2474	-2.015	0.043887 *
<b>Sexo : Faixa Etária</b>					
Sex.Fe:Fx.Et.Um (V.ref.)	28/132 (03%)				
Sex.Fe:Fx.Et.Dois	06/238 (01%)	-1.8027	0.5380	-3.351	0.000806 ***
Sex.Fe:Fx.Et3.Três	14/297 (01%)	1.1827	0.6992	1.692	0.090714.

MOD1 = GLM (CON.NOMINAL ~ SEXO\*FX.ETÁRIA + NÚM.CONST.FLEX.SN + SAL.FÔNICA, DATA = DS, FAMILY = BINOMIAL)

Fonte: Elaborada pela autora, com dados da pesquisa.

No modelo de regressão logística de efeitos fixos, com dados da comunidade Volta do Angico, as estimativas com valor-p significativo sugerem como favorecedores da marcação canônica de plural os fatores: *Faixa Etária* (valor de referência *Faixa D*); *número de constituintes flexionáveis do SN* (valor de referência *Dois constituintes flexionáveis*); *Saliência fônica* (valor de referência *plural irregular*).

Como afirma Scherre (1988, p. 430), a influência de variáveis sociais em fenômenos linguísticos variáveis, principalmente nos casos em que uma das variantes é estigmatizada, não constitui nenhuma novidade dentro da literatura linguística. Estudos, como os já descritos em seções anteriores, nos permitem entender que, em situação de variação que não envolve

mudança, ou seja, na variação estável, normalmente se verifica uma relação proporcional entre as variáveis sociais e uso de formas de prestígio. Por isto, cabe-nos compreender que

não se pode mais seriamente defender que o linguista deve limitar suas explicações da mudança às influências mútuas dos elementos lingüísticos, definidos pela função cognitiva. Nem se pode argumentar em qualquer sentido sério que o sistema lingüístico em mudança é autônomo. [...] não é possível concluir uma análise das relações estruturais dentro de um sistema lingüístico, sem considerar as relações externas (Labov, 1972 [1965], p. 181-182 *apud* Lucchesi, 2004, p. 186).

Essa perspectiva, iniciada por Labov e adotada por outros tantos pesquisadores, possibilita a inserção do componente social na abordagem da mudança linguística e demonstra a importância da inclusão de condicionantes sociolingüísticos, já que:

todos esses estudos demonstraram que a mudança não é apenas uma função do sistema lingüístico, mas uma função da interação da estruturação interna da língua com o processo social em que ela se realiza; sendo que, em muitos casos, os fatores funcionais e estruturais internos podem ser totalmente sobrepujados pela força das disposições sociais (Lucchesi, 2004, p. 185).

Considerando essas proposições, em que se coloca para a investigação da variação e da mudança da língua a necessidade de se correlacionar fatores lingüísticos e socioculturais, temos o fator as *Faixas Etárias II e III* (valor de referência *Faixa I*) com desfavorecedora da marca canônica de plural nos SNs.

A variável faixa etária revela-se de fundamental importância no estudo dos fenômenos variáveis, uma vez que é essa variável que vai indicar se determinado fenômeno de variação linguística está estável ou está em progresso num estudo em tempo aparente. Sabe-se também que, nos eixos sociais, por exemplo, em geral, os falantes mais jovens são mais inovadores e os falantes mais velhos costumam preservar formas mais antigas. O efeito da variável faixa etária sobre a presença de marcas de plural nos SNs pode ser observado, inicialmente, a partir da tabela e gráfico que segue.

**Tabela 7** – Dados comparativos: uso da concordância no SN em função da faixa etária no PB

FAIXA ETÁRIA	SALVADOR -BA (LOPES, 2001) <sup>128</sup>		HELVÉCIA - BA (ANDRADE, 2003) <sup>129</sup>		VOLTA DO ANGICO – BA	
	%	Peso relativo	%	Peso relativo	Estimativa	(p)
Faixa I	81	.45	14	<b>.74</b>	<b>Valor de Referência</b>	
Faixa II	80	<b>.46</b>	9	.48	-0.5999	0.023225 *
Faixa III	85	.59	4	.23	-3.0134	7.31e-07 ***

Fonte: Elaborada pela autora.

Como podemos observar a faixa etária mais nova é tomada como valor de referência, enquanto na Faixa III (valor de estimativa negativo de -3.0134 e valor-p significativo 7.31e-07\*\*\*) temos uma maior favorecimento na não marcação de plural em todos os elementos do SN, seguido da Faixa II (valor de estimativa negativo de -0.5999 e valor-p significativo 0.023225\*). Alguns fatores podem ser aventados diante desse contexto que indica um incipiente processo de mudança em curso, a saber:

- (i) As fichas dos participantes demonstram que mesmo que a amostra se caracterize enquanto representativa do português brasileiro, os falantes mais jovens são os poucos que ainda tiveram acesso à rede educacional (até o 5º ano do Ensino Fundamental), ocasionando a aquisição de marcas em função de um uso mais padrão, sob influências linguísticas de fora da comunidade;
- (ii) A transição periódica campo-cidade dos jovens à procura de trabalho na cidade sede – Canarana – ou na cidade de Irecê, principalmente em época de colheita de cebola e mamona, ampliando sua rede de contatos, enquadrando-se em um continuum de rede social de mais isolada para mais integrada (Bortoni-Ricardo, 2021) (*cf.* Figura 10);

<sup>128</sup> Lopes (2001) trabalha com quatro faixas etárias, Faixa I correspondendo de 15 a 24 anos; Faixa II de 25 a 35 anos; Faixa III de 45 a 55 anos; Faixa IV acima de 65 anos. Portanto consideraremos nesta análise apenas as três últimas faixas etárias. Então onde lê Faixa I entende-se, nos parâmetros de Lopes (2001), por faixa II, onde lê Faixa II entende-se faixa III e onde lê faixa III entende-se faixa IV.

<sup>129</sup> Andrade (2003) trabalha com Faixa I correspondendo de 21 a 40 anos; Faixa II de 41 a 60 anos; Faixa III mais de 60 anos.

- (iii) O processo de democratização da comunicação, em especial a chegada da internet na Comunidade<sup>130</sup> afeta diretamente os mais jovens, que através dos seus smartphones acessam redes sociais e os aplicativos, os quais têm favorecido o acesso a um contexto aproximativo da norma canônica para marcação de plural. Importante deixar claro que a comunidade possui acesso à internet desde 2015, no entanto, o valor mensal pelo acesso foge dos padrões econômicos da Comunidade, o que acarreta o acesso apenas à cerca de 50% dos seus moradores, entre eles, sem dúvidas, os mais jovens.

Portanto, os dados e as hipóteses de possíveis fatores condicionantes apontam que “[...] as primeiras gerações, diante dessa situação, têm propensão a demonstrarem mais relação estreita com essa história de aprendizagem. As gerações posteriores, com o contato com as diversas utilizações da linguagem e, ainda, com o convívio escolar com o padrão, tendem a incorporar, cada vez mais, traços da língua comum, perdendo vestígios da língua inicial” (Baxter, 2001, p. 188-189). Sendo assim, cabe compreendemos que o fator faixa etária pode estar relacionada com outros aspectos sociais, tais como os apontado em (i), (ii) e (iii), escolaridade, mercado de trabalho e rede de relações sociais, permitindo inferir que “[...] nem todo indício de mudança em curso apresentado pela distribuição dos resultados em função da faixa etária é reflexo somente da gradação etária dos falantes que constituem a amostra do estudo” (Freitag, 2005, p.106).

Ainda em relação a *Faixa Etária*, importante relatar que a sincronia entre os dados apresentados por Andrade (2003), para Helvécia, e os da Comunidade Volta do Angico vêm ratificando que as comunidades afro-brasileiras são resultados contemporâneos de um processo de contato linguístico, entre o português e as línguas africanas, e as heranças desse contato linguístico podem ser observadas ainda hoje, principalmente, no léxico e na morfologia, no nosso caso em especial, na expressiva redução da morfologia flexional nos nomes.

No tocando as variáveis preditoras socioculturais ainda cabe discutirmos, mesmo que brevemente, sobre a variável Sexo, a qual foi selecionada na análise univariada, após teste qui-quadrado, no entanto, na rodada multivariada não apresenta fatores cujas estimativas tenham valor-p significativo, o que sugere pouca relação da variável com o processo de concordância.

---

<sup>130</sup> Importante relatar que a comunidade possui acesso à internet desde 2015, no então o valor mensal foge dos padrões econômicos da Comunidade, o que acarreta o acesso à apenas cerca de 50% dos seus moradores, entre eles, sem dúvidas, os mais jovens.



Vejam os dados de Helvécia (Andrade, 2003):

**Tabela 8** – Dados comparativos: uso da concordância no SN em função do gênero no PB

SEXO	SALVADOR LOPES (2001)		HELVÉCIA ANDRADE (2003)		VOLTA DO ANGICO	
	%	Peso relativo	%	Peso relativo	Estimativa	(p)
Masculino	80%	.47	10	.63	Valor de Referência	
Feminino	82%	.53	9	.39	-0.1088	0.687534

Fonte: Elaborada pela autora.

Lembrando que esse valor de referência apresentado nos dados de Volta do Angico é só uma tendência, visto que não há estimativa que esteja demonstrando a variante como significativa. Nestas condições estatísticas, os dados nos revelam que assim como no dialeto de Helvécia (.39 para as mulheres e de .63 para os homens) há uma tendência de que as mulheres realizem menos a concordância canônica na marca de plural do que os homens, diferentemente dos dados (.53 para as mulheres e de .47 para os homens) de Lopes (2001), para Salvador, comunidade essencialmente urbana e não afro-brasileira.

O quadro comparativo, mesmo com plataformas estatísticas distintas, demonstra que os dados de Volta do Angico dialogam com os de Andrade (2003) em que há uma tendência de que dos homens realizem mais a concordância do que as mulheres, diferentemente do que demonstrou Lopes (2001) em suas análises. Tal contexto pode ser compreendido pelo fato dos homens, em comunidades rurais afro-brasileiras, ainda serem mais expostos a contextos sociais diversos, processo de comercialização, que os colocam “[...] numa situação de maior exposição aos padrões lingüísticos adventícios e numa posição que favorece mais a sua assimilação” (Lucchesi, 2000, p. 289-90), enquanto que as mulheres ainda se mantêm mais dedicadas as tarefas do lar e à educação dos filhos “[...] e a falta de uma maior exposição ao convívio social acarreta na preservação de formas lingüísticas mais características do dialeto” (Andrade, 2003, p. 131). Sendo assim, os resultados refletem o que geralmente se espera de uma comunidade rural afro-brasileira, visto que

os homens têm um maior grau de integração social, o que implica um maior contato com padrões lingüísticos adventícios. Passam mais tempo fora da comunidade (deslocam-se para outras cidades da região para vender produtos das roças, ou para trabalhar por temporadas e também trabalham em empresas que funcionam na região,

saindo diariamente da comunidade), enquanto as mulheres tendem a estar mais enraizadas na sua própria comunidade, dedicadas à roça e ao trabalho doméstico (ANDRADE, 2003; BAXTER; LUCCHESI; GUIMARÃES, 1997; LUCCHESI, 2000a) (Baxter, 2008, p. 285).

As variáveis socioculturais são caminhos que nos conduzem a possíveis hipóteses do maior uso do morfema de plural pelos falantes, mas não podemos deixar de ressaltar a importância de tais variáveis conjugadas às variáveis linguísticas, tendo sido duas dessas últimas indicadas pela análise estatística como condicionantes da marca canônica de plural, a saber: *número de constituintes flexionáveis do SN* (valor de referência *SNs com 2 itens flexionáveis*) e variável *Saliência fônica* (valor de referência *Plural Irregular*) .

A hipótese aventada em relação ao número de elementos flexionáveis no SN é que sua extensão seja inversamente proporcional à presença de marcas de plural. Retomaremos nossos resultados em comparação com os de Helvécia (Andrade, 2003):

**Tabela 9** – Dados comparativos<sup>131</sup>: uso da concordância no SN em função do número de constituintes flexionáveis

Nº DE CONSTITUINTES FLEXIONÁVEIS DO SN	HELVÉCIA - BA (ANDRADE, 2003)		VOLTA DO ANGICO - BA	
	%	Peso relativo	Estimativa	(p)
<b>SNs com 1 item flexionável</b>				
“ <i>Só quinze dias.</i> ”	15	.66	0.7063	0.003754 **
<b>SNs com 2 itens flexionáveis</b>				
“ <i>Assisti muitos filmes.</i> ”	7	.46	Valor de referência	
<b>SNs com 3 ou mais itens flexionáveis</b>				
“ <i>Esses lugares grandes, né?</i> ”	5	.26	-1.0358	0.031657 *

Fonte: Elaborada pela autora.

Volta do Angico e Helvécia, em mais uma variável, caminham semelhantemente. Vejamos que os SNs que mais favorecem a aplicação da regra de concordância são aqueles constituídos por apenas um elemento flexionável (para Volta do Angico valor de estimativa positiva de 0.7063 e valor-p 0.003754 \*\* e para Helvécia peso relativo .66), sendo os SNs com três ou mais constituintes pluralizáveis um fator que desfavorece a marca canônica de plural

<sup>131</sup> Na pesquisa de Lopes (2001) essa variável não foi controlada, por apresentar apenas as análises mórficas. Por isso, não seguimos a comparação com dados de Salvador, comunidade (+) urbana, com rede mais integrada (Bortoni-Ricardo, 2021), em comparação com as comunidades afro-brasileiras Volta do Antigo e Helvécia, (-) urbanas, com redes sociais mais isoladas (Bortoni-Ricardo, 2021).

flexionável (para Volta do Angico valor de estimativa negativo de -1.0358 e valor-p 0.031657 \*e para Helvécia peso relativo .66). Portanto,

a frequência de SNs mais complexos, com um número maior de constituintes, seriam próprios da fala formal e sobretudo da escrita; sendo, obviamente, uma possibilidade contida na competência linguística do falante, cuja potencialidade pode ser ampliada em condições especiais de uso da língua. Todavia, na fala vernacular, esses SNs desfavorecem a aplicação da regra de concordância, seja porque sobrecarregam o processamento da fala com um grande número de constituintes a serem marcados, seja porque envolvem estruturas menos coesas com adjunção à direita do núcleo nominal. (Lucchesi; Dália, 2022, p. 7376).

Sendo assim, inferimos que quanto maior o número de elementos flexionáveis no SN, a tendência é que o número de marcas de plural decresça, sendo a concordância favorecida nos sintagmas cuja estrutura possua menos constituintes, se considerarmos o princípio da simplicidade, apontado por Lucchesi (2000).

O fator saliência fônica, também selecionado enquanto relevante à análise da concordância nominal de número, embute-se na hipótese de que a quantidade de material existente na oposição entre singular e plural interfere na possibilidade de se fazer a concordância, chegando à conclusão que os itens em que a oposição é mais saliente (mais material fônico perceptível na oposição singular/plural) devem ser os mais marcados com o morfema de plural (Lopes, 2001, p. 194). Importante frisar que

a influência da saliência fônica na presença/ausência de plural nos elementos do SN é medida a partir de uma escala, considerando-se o nível de diferenciação material fônica na oposição singular/plural. Em língua portuguesa, essa oposição pode se dar mediante o acréscimo do morfema de plural em palavras terminadas por vogal oral (*casa/casas*) ou vogal nasal (*algum/alguns; maçã/maçãs*); através da substituição da consoante final por um conjunto formado de vogal mais consoante (*papel/papéis*); pelo acréscimo da terminação -es (*mar/mares; mês/meses*); havendo ditongo nasal a oposição pode ser feita apenas com o acréscimo do morfema de plural (*cidadão/cidadãos; homem/homens*) ou com alterações no grupo vocálico (*pão/pães*); e ainda com mudanças metafônicas (*olho/'ólhos'*), o denominado plural duplo (Andrade, 2003, p. 101, grifo da autora).

Dessa forma, o item lexical não regular tem mais chances de ser marcado, portanto um SN que contenha pelo menos um constituinte não regular apresenta mais chances de conter todas as marcas formais de plural. Vejamos nossos resultados, em comparação os de Helvécia (Andrade, 2003).

**Tabela 10** – Dados comparativos: uso da concordância no SN em função da saliência fônica no PB

SALIÊNCIA FÔNICA	HELVÉCIA - BA (ANDRADE, 2003)		VOLTA DO ANGICO - BA	
	%	Peso Relativo	Estimativa	(p)
SN <sup>s</sup> apenas com itens regulares “Assisti <u>muitos filmes</u> .”	8	.46	-0.4985	0.043887 *
SN <sup>s</sup> contendo itens que apresentam diferenciação fônica quando pluralizados “Ela tinha <u>oito meses</u> .”	15	.66	Valor de referência	

Fonte: Elaborada pela autora.

Ambos *corpora*, de comunidades afro-brasileiras, como esperado, indicam que o princípio da saliência fônica é relevante na aplicação da regra de concordância. Como afirma Andrade (2003, p. 114), outras pesquisas, que analisaram essa variável, apontam que os valores estatísticos sempre apontam para a força condicionadora das formas que se apresentam mais perceptíveis ao falante, assim como demonstrado na *Tabela 10* em que os SNs de itens regulares favorecem a aplicação da regra de concordância (para Volta do Angico valor de referência e para Helvécia peso relativo .46), enquanto os SNs de itens regulares desfavorecem a aplicação da regra de concordância (para Volta do Angico valor de estimativa negativo de -0.4985 e valor-p 0.043887 \*e para Helvécia peso relativo .66).

Considerando a *Tabela 10*, notamos que o quadro de variação na concordância de número no SN no dialeto da Comunidade Volta do Angico, que se assemelha ao de Helvécia, está em conformidade com o que se tem postulado sobre a aquisição da regra a partir de contextos linguísticos mais salientes, demonstrando que formas nominais que apresentam maior diferenciação morfofonológica na oposição singular-plural, a frequência de aplicação da regra é maior do que com formas nominais que não apresentam essa diferenciação morfofonológica para marcar o plural.

Ao tratar das questões implícitas à saliência fônica Lopes (2001) retoma alguns pressupostos de Guy (1981, p. 296), o qual

ao estudar a saliência fônica na concordância na fala carioca, estabeleceu uma relação entre a quantidade de material fônico e a percepção do ouvinte, no momento da aquisição da regra (numa situação de suposta descrioulização), daí considerar os mais facilmente percebidos, ou mais salientes, como os inicialmente aprendidos ou incorporados. As formas em que a oposição singular/plural é mais saliente são, por isso, mais marcadas com o morfema.

Eis o ponto que entra em cena duas situações/opiniões contrárias: Naro (1981) acredita que é a partir dos itens com concordância menos saliente que a variação se inicia; e Guy (1981) acredita que é a partir dos itens com concordância mais saliente que se dá a percepção, na fase de aquisição (Lopes, p. 195). Nesse sentido Andrade (2003, p. 115) acredita que

há toda uma rede de elementos internos e externos envolvida nos processos de mudança, e um desses elementos é a força das formas lingüísticas de maior prestígio na sociedade, portanto, é plenamente razoável que o falante adquira formas da língua culta a partir de contextos lingüísticos mais salientes, por outro lado, não parece concebível a idéia de que o falante possa vir a substituir formas cultas por não cultas em contextos lingüísticos mais perceptíveis. Mas, ao entrever que as formas mais salientes são as que mais favorecem a descrioulização, Guy, de fato, não poderia concordar com a teoria de Naro e Lemle, pois, se há séculos a língua caminha no sentido de perder as marcas de concordância, como determina a sua deriva interna, não seria possível explicar porque os dialetos populares estão em processo de aquisição dessas marcas.

Portanto, nossos dados apresentam indícios estatísticos de que a variação na concordância de número no SN no *corpus* representativo da Comunidade Volta do Angico dialoga com outros estudos sobre a aquisição da regra a partir de contextos lingüísticos mais salientes, além de apresentar as mesmas tendências de outros dialetos populares, a citar o de Helvécia, no sentido de perder as antigas características provenientes de processos de transmissão lingüística irregular.

Findadas as discussões em relação aos resultados alcançados na análise multivariada de efeitos fixos, seguimos para última etapa, e como afirma Oushiro (2021), “mas não menos importante”, a rodada multivariada de efeitos mistos<sup>132</sup>, tendo em vista que “qualquer estudo que tenha efeitos aleatórios deve incluí-los na modelagem estatística” (Oushiro, 2021, p. 307). A autora ainda afirma que o principal resultado a checar num modelo de efeitos mistos é se os mesmos efeitos fixos continuam a ser correlacionados após a inclusão dos efeitos aleatórios – no nosso *corpus* incluímos as variáveis participantes e item lexical. Vejamos as estimativas com a inclusão das variáveis aleatórias supracitadas:

---

<sup>132</sup> Os modelos de efeitos mistos são implementados no R por meio do pacote *lme4*.

**Tabela 11** – Estimativas dos parâmetros do modelo (de regressão logística, modelo linear generalizado de efeitos mistos com interação) da concordância nominal de número – amostra ELHiS

	<b>Apl./N</b>	<b>Estimativa</b>	<b>Erro Padrão</b>	<b>Valor Z</b>	<b>(p)</b>
<b>Sexo</b>					
Masculino (V. ref.)	83/586(7%)				
Feminino	48/667(4%)	-4.113e-01	5.391e-01	-0.763	0.445494
<b>Faixa Etária</b>					
Um (V. ref.)	81/357(7%)				
Dois	33/429(3%)	-1.193e+00	5.305e-01	-2.248	0.024547 *
Três	17/467(1%)	-3.831e+00	1.026e+00	-3.734	0.000189 ***
<b>Número de Constituintes Flexionáveis do SN</b>					
Dois (V. ref.)	95/926(8%)				
Três	05/95(01%)	-2.667e+00	1.054e+00	-2.532	0.011357 *
Um	31/219(2%)	-2.270e-01	5.256e-01	-0.432	0.665810
<b>Saliência Fônica</b>					
Irregular (V. ref.)	28/184(02%)				
Regular	103/1069(8%)	-1.918e+00	8.401e-01	-2.284	0.022397 *
<b>Sexo : Faixa Etária</b>					
Sex.Fe:Fx.Et.Um (V.ref.)	28/132(03%)				
Sex.Fe:Fx.Et.Dois	06/238(01%)	-1.138e+00	8.938e-01	-1.273	0.203036
Sex.Fe:Fx.Et3.Três	14/297(01%)	2.226e+00	1.179e+00	1.889	0.058878

MOD2 = GLMER (CON.NOMINAL ~ SEXO\*FX.ETÁRIA + NÚM.CONST.FLEX.SN + SAL.FÔNICA + (1|INFORMANTE) + (1|ITEM.LÉXICO), DATA = DS, FAMILY = BINOMIAL)

Fonte: Elaborada pela autora com dados da pesquisa

Os resultados do modelo de efeitos mistos, caracterizado pela inclusão de variáveis aleatórias como o item léxico e participante, mostram, como vemos na *Tabela 11*, que a correlação entre Faixa etária, Número de constituintes pluralizáveis e Saliência Fônica se mantém, com exceção do Sexo, como podemos notar no contraste entre o modelo de efeitos fixos (*cf.* Tabela 6) e o modelo de efeitos mistos (*cf.* Tabela 11), exibindo estimativa com valor-p não significativo que mostra o fator desfavorecendo o processo, o que nos leva a crer que as variáveis aleatórias apresentam efeito sobre essa variável, podendo indicar casos de variação intraindividual, por exemplo, um fato que enfatiza a importância “de se realizar uma análise de

efeitos mistos, com a inclusão de efeitos aleatórios, para que se possa confirmar se os efeitos fixos de fato têm influência na variável resposta” (Oushiro, 2021, p. 312-313).

Sem dúvidas, dados sobre o uso variável da concordância nominal de número no PB tem sido recorrente nas últimas décadas em monografias, dissertações e teses, que favorecem um panorama sistemático dessa variação no PB, no entanto, o que potencializa a abordagem nesta tese é a análise contrastiva com Helvécia (Andrade, 2003) e o fato de proporcionar a discussão sobre a formação da realidade sociolinguística brasileira, através da análise de um *corpus* inédito, constituído no âmbito desta pesquisa, que se desdobrou em um Projeto de Pesquisa – Estudos Linguísticos e Históricos do Sertão (ELiHS) – desenvolvido na Universidade do Estado da Bahia (UNEB), Departamento de Ciência Humanas e Tecnologias (DCHT), *Campus XVI* – Irecê. Além do mais, os dados aqui apresentados nos forneceram material empírico para o desenvolvimento do jogo *online* (*cf.* seção 4.3.1) disponível no site do Projeto, que representa uma pequena intervenção pedagógica da pesquisa, com intuito de inspirar outros pesquisadores que trabalhem na perspectiva da sociolinguística quantitativa, mas que também direcionem seus olhares para uma sociolinguística educacional, condessando este mundo estatístico à caminhos que conduzam a uma pedagogia da variação linguística, na tentativa de alcançar uma educação linguística no combate ao preconceito linguístico.

## 6 CONSIDERAÇÕES FINAIS

*“[...] Laus Deo, que Deus seja louvado; porque o trabalho, a meu ver, será muito, de muitos e não será por pouco tempo” (Mattos e Silva, 2008, p. 30).*

Aqui iniciamos nossas considerações finais, com a mesma epígrafe com que abrimos as considerações que iniciaram esta tese, reafirmando despretensiosamente que o que propomos, nesta oportunidade, é, sem dúvida, um trabalho “[...] *de muitos e não será por pouco tempo*”, dentro de um contexto cíclico que avança, se renova e se redimensiona em novos saberes linguísticos.

Ao longo das seções precedentes, buscamos contribuir discursivamente com os estudos do Português Brasileiro, especificamente em sua vertente popular, vista aqui sob uma perspectiva sócio-histórica (Mattos e Silva, 2008), compreendendo, de modo geral, as mudanças provocadas pelo contato linguístico, de forma mais incisiva por parte dos africanos e afrodescendentes. Ainda em perspectiva teórica, apresentamos alguns marcos referentes aos estudos da linguagem, bem como as possibilidades teóricas em torno dos processos que envolvem variação e mudança linguísticas, alcançando a Sociolinguística Laboviana, eixo teórico e metodológico norteador desta pesquisa, que parte do pressuposto de que a variação é inerente às línguas e condicionada tanto por fatores internos quanto externos ao sistema linguístico (Weinreich; Labov; Herzog, 2006 [1968]; Labov, 2008 [1972]).

A contextualização teórica se fez imprescindível, visto que prevíamos que o contexto sócio-histórico da comunidade afro-brasileira Volta do Angico, mesmo não se caracterizando como uma comunidade (-) isolada, nos termos de Helvécia (Andrade, 2003), refletiria condicionamentos no processo de variação da concordância nominal de número, fenômeno analisado nesta tese. Portanto, tomou-se como objetivo verificar se os resultados estatísticos alcançados eram similares aos postos para Helvécia (Andrade, 2003), a partir da hipótese da Transmissão Linguística Irregular (Baxter; Lucchesi, 1997; Lucchesi, 2000).

Expostas as perspectivas teóricas, iniciamos a *Parte II* desta tese, intitulada *Contribuições para a proposta de interiorização linguística*, com destaque para:



(i) **Implementação do Projeto Estudos Linguísticos e Históricos do Sertão (ELiHS)**, que envolve:

- Inédito mapeamento das 148 comunidades afro-brasileiras presentes no Território de Irecê, certificadas pela Fundação Cultural Palmares (FCP) enquanto remanescentes de quilombos (*cf.* seção 4.1.3);
- Inédita descrição sócio-histórica das comunidades afro-brasileiras do Território de Irecê (BA) (*cf.* seção 4.1.4);
- Composição de *corpus* oral inédito, com devido rigor metodológico, da comunidade afro-brasileira Volta do Angico – conforme presidente da associação, com cerca de 200 anos de existência –, situada no município de Canarana (BA);
- Banco de dados sociolinguístico, que reúne 90.317 palavras, disponibilizando documentos editados em diferentes versões e seus metadados;
- Efetiva atuação no tripé – ensino, pesquisa e extensão – que sustenta a Universidade Pública e de qualidade, em especial, do curso de Letras, Língua Portuguesa e Literaturas da Universidade do Estado da Bahia, *campus* XVI – Irecê, ao articular o conhecimento científico ao regional ((*cf.* vídeos divulgados no *site* do projeto em que temos relatos de Monitores; Direção e coordenações da UNEB – Irecê; Coordenação do Mestrado PPGEAFIN; e representantes da Comunidade externa);
- Atuação efetiva de 13 monitores (atualmente 5 bolsistas, incluindo 01 estudante de iniciação científica júnior, residente na comunidade dos Batatas em Ibititá (BA)), colaborando na formação de indivíduos mais conscientes da realidade sociolinguística brasileira e, então, mais qualificados profissionalmente, tendo em vista que estão em um curso de licenciatura em Letras;
- Aprovações em Editais publicados pela Universidade do Estado da Bahia (UNEB):
  - Edital 015/2023 – Projeto de extensão
  - Edital 073/2022 – Projeto de extensão
  - Edital 024/2022 – Financiamento de projetos de extensão
  - Edital 019/2022 – Iniciação Científica Júnior (IC-Júnior)
  - Edital 018/2022 – Iniciação Científica (IC)
  - Edital 042/2021 – Financiamento de projetos de extensão
  - Edital 012/2021 – Projeto de Extensão
- 02 (dois) capítulos de livros publicados (*cf.* seção 4.2);

- 04 (quatro) monografias defendidas (*cf. seção 4.2*).

(ii) **Desenvolvimento do site Projeto Estudos Linguísticos e Históricos do Sertão (ELiHS)**, que envolve:

- Disponibilização, em consonância com os princípios da Ciência Aberta, de acervo oral, os áudios e as transcrições inéditas, com o devido rigor metodológico e sócio-histórico, na rede mundial de computadores, não restringindo as prospecções e resultados de pesquisa a laboratórios físicos, cujo acesso por interessados de maneira geral normalmente é menos simples;
- Direcionamento de holofotes às comunidades que estão “para além do litoral”, que vivem e resistem no Território de Irecê, através da descrição generalizada sobre o contexto sociodemográfico e registros fotográficos.

(iii) **Desenvolvimento do jogo online *ConcordAção***, que envolve:

- Interrelação entre a Sociolinguística Quantitativa e a Sociolinguística Educacional, em uma proposta de compromisso com a Educação, através de uma aplicação didático-pedagógica do resultado desta tese, entendendo, como Ilari (2003, p. 111), que um dos principais desafios da Linguística no século XXI é “mudar a cara do ensino da língua materna”;
- Usuários como protagonistas no processo de ensino-aprendizagem e uma relação direta entre a variação na concordância nominal de número e o uso de tecnologias no contexto da sala de aula – uma proposta que chega neste século, para revolucionar as metodologias de ensino tradicionais, com novas perspectivas educacionais;
- Um jogo desenvolvido cuidadosamente, pensando numa efetiva aproximação entre realidade e virtualidade, no que tange às comunidades do Território;
- Singela contribuição para a didatização da linguística, fazendo chegar às escolas da Educação Básica, especialmente as do campo, situadas nas comunidades, materiais com dados reais de língua, para a abordagem da variação linguística, recomendada por documentos institucionais – Parâmetros curriculares Nacionais (Brasil, 1998) e Base Nacional Comum Curricular (BNCC) (Brasil, 2016); uma contribuição particularmente no que diz respeito à concordância nominal de número;

- A promoção de um ensino identitário e lúdico, inspirando outros trabalhos no campo da Sociolinguística Educacional, em uma rede de pesquisa e produção que visa não apenas ao impacto científico, mas aos impactos educacional e social.

(iv) **Análise estatística da concordância nominal de número em amostra oral da comunidade afro-brasileira Volta do Angico**, que envolve:

- Análise, através da perspectiva teórica da Sociolinguística Variacionista Quantitativa (Labov, 2008 [1972]), sobre a variação na realização de marcas de plural no sintagma nominal em *corpus* oral inédito da comunidade afro-brasileira Volta do Angico (BA), uma amostra representativa do português popular;
- Primeira tese que analisa quantitativamente o fenômeno da concordância nominal de número através do uso da modelagem estatística de regressão logística com resposta binária, a linguagem de programação denominada R (R CORE TEAM, 2020), por meio da interface o *RStudio* (EQUIPE RSTUDIO, 2020);
- Análise contrastiva dos padrões morfossintáticos de concordância comuns à comunidade tradicional afro-brasileira Volta do Angico e à comunidade Helvécia (Andrade, 2003);
- Utilização dos resultados estatísticos para o desenvolvimento do jogo *online ConcordAção*, como suporte pedagógico para o ensino da concordância nominal de número;
- Resultados que apontam, na amostra, que há uma tendência significativa ao uso não canônico na indicação do plural nos SNs, um resultado estatístico aproximativo aos dados apresentados por Andrade (2003) para Helvécia (89,55% para Volta do Angico e 91% para Helvécia), nos permitindo afirmar que, embora a tendência do português brasileiro seja a marcação explícita de número nos elementos que compõem o SN, ainda há comunidades linguísticas, como as comunidades afro-brasileiras Volta do Angico e Helvécia, Bahia, que não conservam a atual variante prototípica, de marcas explícitas, reafirmando o posicionamento de Lucchesi (2009) sobre a ideia da perda de morfologia flexional no PB popular ser um indicador da influência do contato linguístico.
- Resultados que apontam para a força condicionadora das formas que se apresentam mais perceptíveis ao falante, em que os SNs de itens (-) salientes foneticamente desfavorecem a aplicação da regra de concordância (para Volta do Angico, valor

de estimativa negativo de -0.4985 e valor -p 0.043887 \* e, para Helvécia, peso relativo .66). Os dados se comunicam também na variável *número de constituintes flexionáveis*, sendo os SNs com três ou mais constituintes pluralizáveis um fator que desfavorece a marca canônica de plural (para Volta do Angico, valor de estimativa negativo de -1.0358 e valor-p 0.031657 \* e, para Helvécia, peso relativo .66);

- A comunidade de Volta do ainda exhibe marcas não-padrão numa intensidade muito maior do que a exibida em outros dialetos populares, e não encontra paralelo com dialetos que não tenham história similar ao de Helvécia.

Algumas são as possibilidades que se colocam como prospecções futuras, que se veja:

- Ampliação do *corpus* oral do Projeto ELiHS, atendendo a outras comunidades afro-brasileiras, tão representativas quanto Volta do Angico (BA), compreendendo que o trabalho com *corpus* “sem dúvida é um trabalho necessário, como base para a reconstrução do passado lingüístico do português que aqui se formou, o português brasileiro” (Mattos e Silva, 2004, p. 120);
- Ampliação do estudo a outros níveis de análise, bem como a abordagem de outros fenômenos morfossintáticos, com o propósito de obter panorama geral da realidade sociolinguística, de forma específica das comunidades afro-brasileiras do Território de Irecê;
- Continuidade e ampliação do acervo fotográfico, na tentativa de que este recurso, de natureza não-verbal, materialize contextos que envolvam educação, moradia, religião, saúde e lazer das 148 comunidades afro-brasileira do Território;
- Manutenção do *site* do Projeto ELiHS;
- Desenvolvimento de novos jogos *online*, contemplando novos fenômenos lingüísticos, ampliando as possibilidades para o espaço interativo do *site* do Projeto.

Nesse sentido, a tese filiou-se a um viés sócio-histórico, isto é, interpretou-se o fenômeno da variação na concordância de número, de forma contextualizada e contrastiva – com os dados de Helvécia (Andrade, 2003) –, procurando associá-la à melhor compreensão em relação ao processo de formação do português brasileiro, de forma especial em sua variedade popular, a que mais preserva os efeitos do contato entre línguas na história sociolinguística do país. Não podemos deixar de destacar a importância de um maior investimento em pesquisas

que tenham as comunidades afro-brasileiras como protagonistas no cenário científico, delineando, com maior clareza, nosso ainda grande desconhecimento em relação à contribuição africana e afro-brasileira no cenário brasileiro.

Assim, ressaltamos que os resultados e produtos alcançados na finalização desta tese, que se desenvolveu também em parceria interinstitucional (UNEB-UEFS), apresentam potencial de impacto – científico e social – na medida em que colaborarão para a interiorização da descrição linguística, dando visibilidade às comunidades que estão “para além do litoral”, que vivem e resistem no Território de Identidade de Irecê, fazendo-as presentes no mapa da sociolinguística brasileira...

Assim, finalizamos mais uma etapa dessa contínua pesquisa científica, que nos propomos a iniciar no âmbito desta tese, com a fala do grande Guimarães Rosa, em *Grande sertão: veredas* (1979), que diz que no sertão “tudo é e não é [...] o sertão é do tamanho do mundo, o sertão é o sozinho, é o dentro da gente, é o sem lugar [...], mundão de ausências, é uma espera enorme, uma falta que sempre”. Sertão é isso: “de repente, volta a rodear a gente por todos os lados. O sertão é de noite. Sertão é onde manda quem é forte, com as astúcias, [...] o sertão é do tamanho do mundo, o sertão está em toda parte, o sertão é dentro da gente, [...] o sertão é onde o pensamento da gente se forma mais forte do que o lugar” (Bethânia, 2015, p. 157, grifo nosso).

## REFERÊNCIAS

ALBUQUERQUE, W. R. de; FRAGA FILHO, W. *Uma história do negro no Brasil*. Salvador: Centro de Estudos Afro-Orientais; Brasília: Fundação Cultural Palmares, 2006.

ALKMIN, T. Os escravos e a língua: em busca de bases históricas para uma reflexão. In: RAMOS, J.CM.; ALKMIN, M.CA. (org.). *Para a história do Português Brasileiro: estudos sobre mudança linguística e história social*. V. 5. Belo Horizonte: EDUFMG, 2007. p. 465-483.

ANDRADE, P. R. de. *Um fragmento da constituição sócio-histórica do português do Brasil: variação na concordância nominal de número em um dialeto afro-brasileiro*. Dissertação de Mestrado em Língua Portuguesa: UFBA, Salvador, 2003.

ANJOS, R. S. A. dos. *Territorialidade Quilombola: fotos & mapas*. Brasília: Mapas Editora & Consultoria, 2011.

ANJOS, R. S. A. dos. *Territórios das comunidades remanescentes de Antigos A599 Quilombos no Brasil: primeira configuração espacial*. Brasília: Edição do autor, 1999. 92 p.

ARAÚJO, Z. Patrimônio da cultura negra. *Revista Palmares: Cultura afro-brasileira*, Brasília, v. 4, n. 4, p. 5, out. 2008.

AVELAR, J.; GALVES, C.. O papel das línguas africanas na emergência da gramática do português brasileiro. *Linguística* [online]. v. 30, n.2, 2014, p.241-288. Disponível em: <http://www.scielo.edu.uy/pdf/ling/v30n2/v30n2a10.pdf>. Acesso em: 20 abr. 2022.

BARBOSA, E. S. “Conferência de Berlim (1884-1885)”. In: MAGNOLI, D. *História da Paz*. São Paulo: Contexto, 2008, p. 153-176.

BAXTER, A. N. Creole-like features in the verb system of na afro-brasilian variety of portuguese. In: SPEARS, A. K.; WINFORD, D. *The structure and status of pidgin and creoles*. Amsterdam/Philadelphia: John Benjams Publishing Company, 1997. p. 265-288.

BAXTER, A. N. Transmissão geracional irregular na história do português brasileiro: divergências nas vertentes afro-brasileiro. *Revista O foco: o português no mundo*. n. 14, dezembro de 1995. p. 72-90.

BAXTER, A. N. The development of variable NP plural agreement in a restructured african variety of portuguese. In: GENEVIÈVE, E.; ARMIN, S. *Creoles, contact, and language change: linguistic and social implications*, v. 27. Amsterdam: John Benjamins, 2004. p. 97-126.

BAXTER, A. N. Concordância nominal. In: LUCCHESI, D.; BAXTER, A.; RIBEIRO, I. (org.). *O Português Afro-Brasileiro*. Salvador: EDUFBA, 2009.

BAXTER, A. N.; LUCCHESI, D. A relevância dos processos de pidginização e criouliização na formação da língua portuguesa no Brasil. *Estudos Lingüísticos e Literários*, Salvador, n. 19, p. 65-84, 1997.

BECHARA, E.. *Moderna gramática portuguesa*. 37. ed. ver. e ampl. 16. Reimpr. Rio de Janeiro: Nova Fronteira, 2009.

BERBER SARDINHA, T. Linguística de corpus: histórico e problemática. *D.E.L.T.A.*, v. 16, n. 2, 2000, p. 323-367. Disponível em: <https://www.scielo.br/j/delta/a/vGknQkZQGsgYbrQfKmtZY4s/?lang=pt>. Acesso em 29 jun. 2023.

BETHÂNIA, Maria (org.). *Caderno de Poesias*. Belo Horizonte: UFMG, 2015.

Boletim Informativo, NUER (Núcleo de Estudos sobre Identidade e Relações Interétnicas). Fundação Cultural Palmares, v. 1, n. L2 . Florianópolis: UFSC, 199. Coordenação de Iika Boaventura Leite. Disponível em: [https://nuer.ufsc.br/files/2014/04/m-mgpefhgnff\\_boletim\\_nuer\\_\\_1.pdf](https://nuer.ufsc.br/files/2014/04/m-mgpefhgnff_boletim_nuer__1.pdf). Acesso em: 12 fev. 2023.

BONVINI, E. Línguas africanas e português falado no Brasil. In: PETTER, M.; FIORIN, J. L. *África no Brasil: a formação da língua portuguesa*. 1. ed., 2. Reimp. São Paulo: Contexto, 2009.

BONVINI, E; PETTER, M. M. T. *Portugais du Brésil et langues africaines*. *Langages*. Paris, n. 130, 1998.

BORGES, P. de S. *Línguas africanas e português brasileiro: análise historiográfica de fontes e métodos de estudo no Brasil (séc. XIX e XX)*. São Paulo, 2015. Dissertação (mestrado em Linguística) – FFLECH-USP. Disponível em: [https://www.teses.usp.br/teses/disponiveis/8/8139/tde-06082015-114127/publico/2015\\_PatriciaDeSouzaBorges\\_VCorr.pdf](https://www.teses.usp.br/teses/disponiveis/8/8139/tde-06082015-114127/publico/2015_PatriciaDeSouzaBorges_VCorr.pdf). Acesso em: 11 jul. 2023.

BORGES NETO, J. História e Filosofia da Linguística: uma entrevista com José Borges Neto. *ReVEL: Revista virtual de estudos da linguagem*. v. 8, n. 14, 2010. Disponível em: [http://www.revel.inf.br/files/entrevistas/revel\\_14\\_entrevista\\_borges\\_netto.pdf](http://www.revel.inf.br/files/entrevistas/revel_14_entrevista_borges_netto.pdf). Acesso em: 02. ago. 2022.

BORTONI-RICARDO, S. M. *Do campo para a cidade: estudos sociolinguísticos de migração e redes sociais*. São Paulo: Parábola, 2011 [1985].

BORTONI-RICARDO, S. M. *Educação em língua materna: a sociolinguística na sala de aula*. São Paulo: Parábola Editorial, 2004.

BORTONI-RICARDO, S. M. *Nós chegemu na escola, e agora?: sociolinguística & educação*. São Paulo: Parábola Editorial, 2005.

BRAGA, M. L. *A concordância de número no sintagma nominal no triângulo mineiro*. Dissertação de Mestrado. Pontifícia Universidade Católica. Rio de Janeiro, 1977.

BRAGA, M. L.; SCHERRE, M. M. P.. A concordância de número no SN na área urbana do Rio de Janeiro. *Anais do Encontro Nacional de Linguística*. Rio de Janeiro: PUC, 1976. pp. 446-477.

BRANDÃO, S. F. Concordância nominal em duas variedades do português: convergências e divergências. *Revista em Estudos Linguísticos Veredas*. Atemática 1/2011. p. 164-178

BRASIL. Secretaria de Educação Fundamental. *Parâmetros curriculares Nacionais: terceiro e quarto ciclos do Ensino Fundamental: Língua Portuguesa*. Brasília. MEC/SEF, 1998.

BRASIL. Ministério da Educação. *Base Nacional Comum Curricular*. Proposta preliminar. Segunda versão revista. Brasília: MEC, 2016.

BRASIL. *Resolução do Conselho Nacional da Saúde nº 510*, de 07 de abril de 2016. Diário Oficial [da] República Federativa do Brasil, Poder Executivo, Brasília, DF, 24 maio 2016. Seção 1, p. 44-46

BRASIL. *Resolução nº 466*, de 12 de dezembro de 2012. Disponível em: [https://bvsm.sau.gov.br/bvs/sau/legis/cns/2013/res0466\\_12\\_12\\_2012.html](https://bvsm.sau.gov.br/bvs/sau/legis/cns/2013/res0466_12_12_2012.html) Acesso em: 01 maio 2021.

BRASIL. Ministério da Saúde. Conselho Nacional de Saúde. *Manual operacional para comitês de ética em pesquisa*. Brasília: Ministério da Saúde, 2002. Disponível em: [https://conselho.sau.gov.br/biblioteca/livros/Manual\\_ceps.pdf](https://conselho.sau.gov.br/biblioteca/livros/Manual_ceps.pdf). Acesso em: 01 maio 2021.

BRASIL. Resolução CNE/CEB n. 8, de 20 de novembro de 2012. *Define as Diretrizes Curriculares Nacionais para a Educação Escolar Quilombola na Educação Básica*. Diário Oficial da União, Brasília, DF: MEC/CNE/CEB, 21 nov. 2012. Seção 1, p. 26.

BRASIL, 1988. Constituição Federal. Disponível em: <http://www.mds.gov.br/aceso-ainformacao/legislacao/segurancaalimentar/leis/1988/Lei,P20,P20Os,P20Quilombolas,P20na,P20Constituicao,P20Federal,P20de,P201988.pdf.pagespeed.ce.HyAQRTEBv8.pdf> Acesso em: 20 de abr. 2022

BOUQUET, S. *Introdução à leitura de Saussure*. Trad. de Carlos A. L. Salum e Ana Lúcia Franco. São Paulo: Cultrix, 2000 [1997].

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

CÂMARA Jr., J. M. *Estrutura da Língua Portuguesa*. 13. ed. Petrópolis: Editora Vozes, 2000.

CÂMARA Jr., J. Ele como um acusativo no português do Brasil. In: CÂMARA Jr., J. *Dispersos*. Rio de Janeiro: FGV, 1972.

CARDOSO, S. A. M. da S. *Geolinguística: tradição e modernidade*. São Paulo: Parábola Editorial, 2010.

CARNEIRO, Z. de O. N.; ALMEIDA, N. L. F. de. Elementos para uma sócio-história do semi-árido baiano. In: RAMOS, J. M.; ALKMIM, M. A. (org.). *Para a História do Português Brasileiro: estudos sobre mudança linguística e história social*. 1. ed. Belo Horizonte: Faculdade de Letras da UFMG, 2007, v. 1, p. 423-442.



- CARNEIRO, Z. de O. N.; ALMEIDA, N. L. F. de. A criação de escolas a partir de critérios demográficos na Bahia do século XIX: uma viagem ao interior. In: Lobo, T.; RIBEIRO, I.; CARNEIRO, Z.; ALMEIDA, No. (org.). *Para a História do português brasileiro: novos dados, novas análises*. Salvador: Edufba, 2006. v. 2, p. 649-674.
- CARVALHO, C. de. Para compreender Saussure. Rio de Janeiro: Editora Rio, 2003 [1976].
- CASTILHO, A. T. de. Funcionalismo e gramática do português brasileiro. In: SOUSA, E. R. de et. al. *Funcionalismo linguístico: novas tendências teóricas*. São Paulo: Contexto, 2012.
- CASTRO, Y. P. de. O português do Brasil, uma intromissão nessa história. In: GALVES, Charlotte; GARMES, H.; RIBEIRO, F. R. *África-Brasil: caminhos da língua portuguesa*. Campinas, SP: EDUNICAMP, 2009. p. 175-184.
- CASTRO, Y. P. de. *Falares africanos na Bahia: um vocabulário afro-brasileiro*. Rio de Janeiro: ABL/TOPBOOKS, 2005. p. 129.
- COELHO, F. A.. Os dialetos românicos ou neolatinos na África, Ásia e América. In: BARBOSA, J. M. (org.) *Estudos Linguísticos Crioulos*. Lisboa: Academia Internacional de Cultura Portuguesa, 1986 [1967].
- COAN, M. ; FREITAG, R. M. K. . Sociolinguística variacionista: a: pressupostos teórico-metodológicos e propostas de ensino. *Revista Domínios de Lingu@Gem*, v. 4, p. 173-194, 2011. Disponível em: <http://www.seer.ufu.br/index.php/dominiosdelinguagem>. Acesso em 10 jan. 2021.
- COMPANHIA, C. A. *Uso da morfologia flexional de número em sintagmas nominais plurais na interlíngua de crianças da 3ª classe*. Maputo: Universidade Eduardo Mondlane. Tese de doutorado, 2001.
- CUNHA JUNIOR, H. A. História dos afrodescendentes: disciplina do curso de Pedagogia da Universidade Federal do Ceará. *Revista Espaço Acadêmico*, n 232, jan./fev., 2022. Disponível em: <https://periodicos.uem.br/ojs/index.php/EspacoAcademico/article/view/57807>. Acesso em: 12 fev. 2022.
- DINAH, C. De Fernão de Oliveira e da (Socio)linguística. In: ABAURRE, M. B.; PFEIFFER, C. AVELAR, J. (orgs). *Fernão de Oliveira: um gramático na história*. São Paulo: Pontes Editores, 2009.
- DUBOIS, J. et. al.. *Dicionário de linguística*. Trad. Izidoro Blikstein et all. São Paulo: Cultrix, 1973.
- ECKERT, P. Three waves of variation study: the emergence of meaning in the study of sociolinguistic variation. *Annual Review of Anthropology*, 41, Palo Alto, 2012, p. 87-100. Disponível em: <https://web.stanford.edu/~eckert/PDF/ThreeWaves.pdf>. Acesso em: 20 out. 2019.
- ECKERT, P.. *Jocks and Burnouts: Social Categories and Identity in the High School*. New York: Teach. Coll. Press, 1989.

ECKERT, P. 2000. *Linguistic Variation as Social Practice*. Oxford: Blackwell

FAGE, J. D. *História da África*. Lisboa: Edições 70, 2010.

FARACO, C. A. *Norma culta brasileira: desatando alguns nós*. São Paulo: Parábola Editorial, 2008.

FARACO, C. A. *Linguística História: uma introdução ao estudo da história das línguas*. São Paulo: Parábola, 2005 [1950].

FARACO, C. A.; ZILLES, A. M. S. (org.). *Pedagogia da variação linguística: língua diversidade e ensino*. São Paulo: Parábola Editorial, 2015.

FERREIRA, E. A. 2019. *Políticas de desenvolvimento territorial e cultural no território de identidade de Irecê-BA: uma análise do investimento em ações e projetos culturais através do programa fundo de cultura no período de 2009 a 2017*. Monografia do curso de Economia da Universidade Federal da Bahia, 2019. Disponível em: <https://repositorio.ufba.br/bitstream/ri/31834/1/TCC2%20ELQUER.docx%20COMPLETO%201.pdf>. Acesso em: 15 maio 2022.

FERREIRA, F. A. *No festim dos lobos: o império do Brasil e o colonialismo na África Centro-Occidental (1880-1885)*. 2018. 462 f. Tese (Doutorado em História) - Instituto de Ciências Humanas e Sociais, Universidade Federal Rural do Rio de Janeiro, Seropédica, 2018. Disponível em: <https://tede.ufrj.br/jspui/handle/jspui/4747>. Acesso em: 9 fev. 2023.

FIGUEIREDO, Carlos Filipe Guimarães. A concordância variável no sintagma nominal plural do português reestruturado de Almojarife. *PAPIA* 18. 23-43, 2008.

FREITAG, R. M. Ko. Banco de dados falares sergipanos: falares sergipanos database. *Working Papers em Linguística*, v. 14, Florianópolis, abr./jul, 2013, p. 156-164. Disponível em <http://dx.doi.org/10.5007/1984-8420.2013v14n2p156>. Acesso em: 16 de jun. de 2021.

FREITAG, R. M. Ko *et al.* Desafios da gestão de dados linguísticos e a ciência aberta. *Cadernos de Linguística*. v. 2, n. 1, 2021, p. 01-19. Disponível em: <https://cadernos.abralin.org/index.php/cadernos/article/view/307>. Acesso em: 30 abr. 2021.

FREITAG, R. M. Ko. Amostras sociolinguísticas: probabilísticas ou por conveniência? *Revista de Estudos da Linguagem*, v. 26, n. 2, p. 667-686, 2018. Disponível em: <http://www.periodicos.letras.ufmg.br/index.php/relin/article/view/12412>. Acesso em 01 de jul. de 2021.

FREITAG, R. M. Ko. *Documentação sociolinguística [recurso eletrônico]: coleta de dados e ética em pesquisa*. São Cristóvão: Editora UFS, 2017. Disponível em: [https://www.researchgate.net/publication/319186082\\_Documentacao\\_sociolinguistica\\_coleta\\_de\\_dados\\_e\\_etica\\_em\\_pesquisa](https://www.researchgate.net/publication/319186082_Documentacao_sociolinguistica_coleta_de_dados_e_etica_em_pesquisa). Acesso em: 26 maio 2021.

FREITAG, R. M. Ko; MARTINS, M. A.; TAVARES, M. A. Bancos de dados sociolinguísticos do português brasileiro e os estudos de terceira onda: potencialidades e limitações. *Alfa Rev. Linguíst.* (São José Rio Preto) [online]. 2012, vol.56, n.3, p.917-944.

Disponível em: <https://doi.org/10.1590/S1981-57942012000300009>. Acesso em: 30 abr. 2021.

GUSMÃO, N. M. M. de. Direitos específicos: vazio legal e luta étnica. *XVII Encontro Anual da Associação de Pós-Graduação e Pesquisa em Ciências Sociais – ANPOCS*, 18 pág. UNESP, 1993. Disponível em: <https://www.anpocs.com/index.php/encontros/papers/18-encontro-anual-da-anpocs/gt-17/gt15-13/7471-neusagusmao-identidade/file>. Acesso em: 15 abr. 2023.

GUSMÃO, N. M. M. de. Negro e Camponês: política e identidade no meio rural brasileiro. *Revista da Fundação SEADE*, 6 (3). S. Paulo, Jul.1./Set. de 1992. Disponível em: [http://produtos.seade.gov.br/produtos/spp/v06n03/v06n03\\_15.pdf](http://produtos.seade.gov.br/produtos/spp/v06n03/v06n03_15.pdf). Acesso em: 12 fev. 2023.

GUY, G. R. Introdução à análise quantitativa da variação linguística. In: GUY, G. R.; ZILLES, A. *Sociolinguística quantitativa: instrumental de análise*. São Paulo: Parábola Editorial, 2007. p. 19-46.

GUY, G. R. A questão da crioulização no português do Brasil. In: ZILLES, A. M. S. (org.). *Estudos de variação linguística no Brasil e no Cone Sul*. Porto Alegre: Editora da UFRGS, 2005.

GUY, G. *Linguistic variation in brasilian portuguese: aspects of phonology, syntax and language history*. 1981. 391 f. Tese (Doutorado em Linguística) Faculdade da Universidade de Pensilvania, Pensilvania.1981.

GUY, G. R.; ZILLES, A. *Sociolinguística quantitativa: instrumental de análise*. São Paulo: Parábola Editorial, 2007.

IBGE. *Instituto Brasileiro de Geografia e Estatística*. Disponível em: <https://www.ibge.gov.br/pt/inicio.html>. Acesso em: 10 de maio de 2021.

ILARI, R. Rodolfo. In: XAVIER, A. C; CORTEZ, S. (org.). *Conversas com linguistas: virtudes e controvérsias da linguística*. São Paulo: Parábola, 2003. p. 97-112.

INCRA. *Instituto Nacional de colonização e Reforma Agrária*. Disponível em: <https://antigo.incra.gov.br/pt/>. Acesso em: 01 de jul. de 2021.

INVERNO, L. A transição de Angola para o português vernáculo: estudo morfossintático do sintagma nominal. In: CARVALHO, Ana Maria. *Português em Contacto*. Madrid-Frankfurt: Vervuer, 2009.

INVERNO, L. *Angola's transition to vernacular portuguese*. Coimbra: Universidade de Coimbra, Faculdade de Letras, Dissertação de Mestrado, 2005.

INVERNO, L. Português vernáculo do Brasil e português vernáculo de Angola: Reestruturação parcial vs. mudança linguística. In: FERNÁNDEZ-FERREIRO, M.; VEIGA, N. V. *Los criollos de base ibérica*: ACBLPE. Madrid: Iberoamericana, 2004.

HALL, Stuart. *Da diáspora: identidades e mediações culturais*. Belo Horizonte: Ed. UFMG; Brasília: UNESCO, 2003.

HOLM, J. Popular Brazilian Portuguese: a semi-creole. In: D'ANDRADE, E.; KIHN, A. (Org.). *Actas do colóquio sobre crioulos de base lexical portuguesa*. Lisboa: Edições Colibri, 1992. pp. 37-66.

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. pp 406-429.

HOUAISS, A. *O português no Brasil*. Rio de Janeiro: Unibrade, 1985.

JACOBETTY, Pedro. *Ciência aberta: produção do conhecimento científico na sociedade em rede*. 2010. Dissertação (Mestrado em Sociologia). Instituto Universitário de Lisboa. Lisboa, 2010. Disponível em: <https://repositorio.iscte-iul.pt/handle/10071/3024>. Acesso em: 28 jul. 2023.

JEROSLOW, E. H. M. *Rural Cearense Portuguese: a study of one variety of nonstandard Brazilian speech*. (Dissertation). Cornell University, 1974.

JOHNSON, D. E. Getting off the GoldVarb Standard: Introducing Rbrul for Mixed-Effects Variable Rule Analysis. *Language and Linguistics Compass* 3/1, 2009. p. 359–383.

JON-AND, A. Concordância variável de número no SN do português L2 de Moçambique: algumas explicações sociais e linguísticas. *Revista de Crioulos de Base Lexical Portuguesa e Espanhola* 2. 2010. p. 28-50.

LABOV, W. *Padrões sociolinguísticos*. Tradução Marcos Bagno, Maria Marta Pereira Scherre e Caroline R. Cardoso. São Paulo: Editora Parábola, 2008 [1972].

LABOV, W. *Principles of linguistic change*. Malden/Oxford: Blackwell, 1994.

LABOV, W. *The Social Stratification of English in New York City*. Washington, DC: Cent. Appl. Ling, 1966.

LABOV, W. *The social motivation of a sound change*. *Word* 18:1–42, 1963.

LEITE, I. B. Os quilombos no Brasil: questões conceituais e normativas. *Etnográfica*, v. 4, n. 2, p. 333-354, 2000. Disponível em < [http://ceas.iscte.pt/etnografica/docs/vol\\_04/N2/Vol\\_iv\\_N2\\_333-354.pdf](http://ceas.iscte.pt/etnografica/docs/vol_04/N2/Vol_iv_N2_333-354.pdf) >. Acesso em 14 de agosto de 2021.

LEMLE, M.; NARO, A. J. *Competências básicas do português*. Relatório final de pesquisa apresentado às instituições patrocinadoras da fundação Movimento Brasileiro de Alfabetização (MOBRAL) e Fundação Ford. Rio de Janeiro: Fundação Movimento Brasileiro de Alfabetização, 1977.

LEMES, E. de Q. *Biometria de sementes, morfologia e crescimento inicial de cinco espécies arbóreas*. 2011. Dissertação de mestrado do Programa de Pós-graduação em Ciências Agrárias da Universidade Federal do Espírito Santo. 2011.

LOBO, T. Rosa Virgínia Mattos e Silva e a história social linguística do Brasil. *Estudos de Lingüística Galega*, v. 7, 2015, p. 69-82

LOBO, T. A questão da periodização da história linguística do Brasil. In: CASTRO, I.; DUARTE, I. *Razões e emoção: miscelânea de estudos em homenagem a Maria Helena Mateus*. Lisboa: Imprensa Nacional - Casa da Moeda, 2003. p. 395-410

LOPES, N. da S. *Concordância nominal, contexto linguístico e sociedade*. Tese de Doutorado. Universidade Federal da Bahia – Instituto de Letras, Salvador. 2001.

LOURENÇO, E. *A nau de Ícaro e imagem e miragem da lusofonia*. São Paulo: Companhia das Letras, 2001.

LUCCHESI, D. A periodização da história sociolinguística do Brasil. *DELTA* [online], v. 33, n. 2, p. 347-382, 2017. Disponível em: <http://dx.doi.org/10.1590/0102-445067529349614964>. Acesso em: 20 maio 2023.

LUCCHESI, D. *Língua e sociedades partidas: a polarização sociolinguística no Brasil*. São Paulo: Contexto, 2015.

LUCCHESI, D. A deriva secular na formação do português brasileiro: uma visão crítica. In LOBO, T. et al. (org). *Rosae: linguística histórica, história das línguas e outras histórias* [online]. Salvador: EDUFBA, 2012, pp. 249-274.

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. Aspectos gramaticais do português brasileiro afetados pelo contato entre línguas: uma visão de conjunto. In: RONCARATI, C.; ABRAÇADO, J. (org.) *Português Brasileiro II: contato linguístico, heterogeneidade e história*, 366-390. Niterói: EDUFF, 2008.

LUCCHESI, Dante. Parâmetros sociolinguísticos do português brasileiro. *Revista da ABRALIN*, v. 5, n. 1 e 2, 2006. p. 83-112.

LUCCHESI, D. A sistematicidade da mudança: do Estruturalismo à Sociolinguística. In: LUCCHESI, D. *Sistema, mudança e linguagem: um percurso na história da linguística moderna*. São Paulo: Parábola, 2004. p. 165-180.

LUCCHESI, D. O conceito de transmissão linguística irregular e o processo de formação do português do Brasil. In: RONCARATI, C.; ABRAÇADO, J. (org.). *Português brasileiro: contato linguístico, heterogeneidade e história*. Rio de Janeiro: 7Letras, 2003. p. 272-284.

LUCCHESI, D. Norma linguística e realidade social. In: BAGNO, M. (org.). *Linguística da norma*. São Paulo: Edições Loyola, 2002. p. 63-92

LUCCHESI, D. As duas grandes vertentes da história sociolinguística do Brasil, *D.E.L.T.A.*, São Paulo, n. 17, v. 1, p. 97-130, 2001.

LUCCHESI, D. Variação e norma: elementos para uma caracterização sociolinguística do português do Brasil. *Revista Internacional de Língua Portuguesa*, n.12, p. 17-28, 1994.

LUCCHESI, Dante. A constituição histórica do português brasileiro como um processo bipolarizador: tendências atuais de mudança nas normas culta e popular. In: GROBE, S.; ZIMMERMANN, K. (org.). *'Substandard' e mudança no português do Brasil*. Frankfurt: TFM, 1998.

LUCCHESI, D.; BAXTER, A. A transmissão linguística irregular. In: Dante Lucchesi, Alan Baxter, Ilza Ribeiro (Org.). *O português afro-brasileiro*. Salvador: EDUFBA, 2009. p. 101-124.

LUCCHESI, D.; DÁLIA, J. Novos condicionamentos estruturais da variação na concordância nominal de número. *Forum lingüístic.*, Florianópolis, v.19, n.1, p. 7369 - 7386, jan./mar. 2022. Disponível em: <https://dialnet.unirioja.es/servlet/articulo?codigo=8491593>. Acesso em 23 jul. 2023.

MARQUES, I. G. Algumas considerações sobre a problemática linguística em Angola. In: Instituto de Cultura e Língua Portuguesas (ed.). *Actas do congresso sobre a situação actual da língua portuguesa no mundo*. Lisboa: Instituto de Cultura e Língua Portuguesas, 1985.

MARQUES, M. G.; LACERDA, M. F. O. *Tour virtual pelo CE-DOHS: Cartas para Vários Destinatários*. Disponível em: <https://youtube.com/playlist?list=PLj5rJKWcYZxNpFsuGwGEmuauc2cJTqocS>. Acesso em: 14 fev. 2023.

MARQUILHAS, R.; HENDRICKX, I. Avanços nas humanidades digitais. In: MARTINS, A. M.; CARRILHO, E. *Manual de Linguística Portuguesa*. MRL Series. De Gruyter, 2016. p. 1-26.

MARTELOTTA, M. E. Linguística. In: MARTELOTTA, M. E et al. (org.). *Manual de Linguística*. 2. ed. São Paulo: Contexto, 2011.

MARTELOTTA, M. E.; KENEDY, E. A visão funcionalista da linguagem no século XX. In: CUNHA, M. A. F. da; OLIVEIRA, M. R. de; MARTELOTTA, M. E. *Linguística funcional: teoria e prática*. Rio de Janeiro: FAPERJ: DP&A, 2003, p. 17-29.

MATTOS E SILVA, R. V. Para a história do português culto e popular brasileiro: sugestões para uma pauta de pesquisa. *Cadernos de Letras da UFF – Dossiê: Literatura, língua e identidade*. N. 34, p. 11-30, 2008a.

MATTOS E SILVA, R. V. Teorias da mudança linguística e a sua relação com a(s) história(s) da(s) língua(s). *Revista de Estudos Linguísticos da Universidade do Porto* - v. 3 – 2008b. p. 39-53.

MATTOS E SILVA, R. V. *Ensaio para uma sócio-história do Português Brasileiro*. São Paulo: Parábola Editorial, 2004a.

MATTOS E SILVA, R. V. *O português são dois: novas fronteiras, velhos problemas*. São Paulo: Parábola, 2004b.

MATTOS E SILVA, R. V. Para a história do português culto e popular brasileiro: sugestões para uma pauta de pesquisa. In: ALKMIM, T. M. (Org.). *Para a história do português brasileiro*. v. 1. São Paulo: Humanitas/FFLCH, 2002. p. 443-464.

MATTOS E SILVA, R. V.. De fontes sócio-históricas para a história social linguística do Brasil: em busca de indícios. In: MATTOS E SILVA, R. V. (Org.). *Para a história do português brasileiro: primeiros estudos*. São Paulo: Humanitas/FFCHL/USP; FAPESP, v. 2, t. 2, 2001. p. 275-302.

MATTOS E SILVA, R. V. Variação, mudança e norma (movimentos no interior do português brasileiro). In: CARDOSO, S. A. M. (Org.). *Diversidade linguística e ensino*. Salvador: EdUFBA, 1996. p.19-43. Disponível em: [https://www.prohpor.org/\\_files/ugd/c8e334\\_843e0a32c1064f1a82304dd3d41611d3.pdf](https://www.prohpor.org/_files/ugd/c8e334_843e0a32c1064f1a82304dd3d41611d3.pdf). Acesso em: 02 ago. 2022.

MATTOS E SILVA, R. V. *Contradições no ensino de português: a língua que se fala X a língua que se ensina*. São Paulo: Contexto, 1996

MELLO, H. R. Português padrão, português não-padrão e a hipótese do contato linguístico. In: ALKMIM, T. M. (Org.). *Para a história do português brasileiro: Novos estudos*. v. III. São Paulo: Humanitas, 2002. p. 341-358.

MINGAS, A. A. *Interferência do Kimbundo no Português falado em Luanda*. Luanda: Caxinde, 2000.

MOLLICA, C. Fundamentação teórica: conceituação e delimitação. In: MOLLICA, Cecília; BRAGA, M. L. (org.). *Introdução à Sociolinguística: o tratamento da variação*. São Paulo: Contexto, 2004, p. 9-14.

MOLLICA, C.; BRAGA, M. L. *Introdução à sociolinguística: o tratamento da variação*. 3. ed. São Paulo: Contexto, 2008.

MONTEIRO, K. dos S.; GARCIA, M. F. Dos territórios de reforma agrária à territorialização quilombola: o caso da comunidade negra de Gurugi, Paraíba. *Revista Pegada*, v. 11, n. 2, 2010. Disponível em: <https://revista.fct.unesp.br/index.php/pegada/article/viewFile/1310/1310>. Acessado em: 14 de ago. de 2021.

MUSSA, A. *O papel das línguas africanas na história do português no Brasil*. Dissertação de mestrado. Rio de Janeiro: UFRJ, 1991, mimeo.

NASCIMENTO, M. B. Historiografia do Quilombo. In: NASCIMENTO, M. B. *Beatriz Nascimento, Quilombola e intelectual: possibilidade nos dias da destruição*. Editora Filhos a África, 2018a, p. 273-294. Disponível em: [https://edisciplinas.usp.br/pluginfile.php/4934266/mod\\_resource/content/1/Untitled\\_29082019\\_193614.pdf](https://edisciplinas.usp.br/pluginfile.php/4934266/mod_resource/content/1/Untitled_29082019_193614.pdf). Acesso em: 06 jun. 2022.

NASCIMENTO, M. B. O conceito de quilombo e a resistência cultural negra. In: NASCIMENTO, M. B. *Beatriz Nascimento, Quilombola e intelectual: possibilidade nos dias da destruição*. Editora Filhos a África, 2018b, p. 273-294. Disponível em:

[https://edisciplinas.usp.br/pluginfile.php/4934276/mod\\_resource/content/1/Untitled\\_29082019\\_194415.pdf](https://edisciplinas.usp.br/pluginfile.php/4934276/mod_resource/content/1/Untitled_29082019_194415.pdf). Acesso em: 06 jun. 2022.

NARO, A. J. O conceito de transmissão linguística irregular e as origens estruturais do português brasileiro: um tema em debate. In: RONCARATI, C.; ABRAÇADO, J. *Português brasileiro: contato linguístico, heterogeneidade e história*. Rio de Janeiro: 7Letras, 2003. p. 285-302.

NARO, A. J.; SCHERRE, M. M. P. *Origens do português brasileiro*. São Paulo: Parábola Editorial, 2007.

NARO, A. J.; SCHERRE, M. O conceito de transmissão linguística irregular e as origens estruturais do português brasileiro: um tema em debate. In: RONCARATI, C.; ABRAÇADO, J. (org.). *Português brasileiro: contato linguístico, heterogeneidade e história*. Rio de Janeiro: 7 Letras, 2003.

NARO, A. J.; SCHERRE, M. Sobre as origens do português popular do Brasil. *DELTA*, v. 9, n. Especial, 1993, p. 437-454.

A NARO, A. J.; SCHERRE, M.. Variação e mudança linguística: fluxos e contrafluxos na comunidade de fala. *Cadernos de Estudos Linguísticos* (20): 9-16). Campinas: Unicamp, Jan./Jun. 1991.

NEGRÃO, E. V.; VIOTTI, E. Em busca de uma história linguística. *Rev. Est. Ling.* Belo Horizonte, v. 20, n. 2, jul./dez. 2012. p. 309-342.

NEVES, M. H. de M. *A gramática funcional*. São Paulo: Martins Fontes, 1997.

NOLL, V. *O português brasileiro: formação e contrastes*. São Paulo: Globo, 2008

OLIVEIRA, J. M. *O futuro da língua portuguesa ontem e hoje: variação e mudança*. 2006. 254f. Tese de doutorado. Faculdade de Letras, Universidade do Rio de Janeiro, Rio de Janeiro.

OLIVEIRA, K.; SOLEDADE, J.; SANTOS, V. de S. Concordância nominal. In: OLIVEIRA, K.; LOBO, T. (org.). *África à vista: dez estudos sobre o português escrito por africanos no Brasil do século XIX*. Salvador: Edufba, 2009. pp. 255-316.

OUSHIRO, L. Introdução à Estatística para Linguistas. *Revista Abralín*, 2022. Disponível em: <https://editora.abralin.org/wp/wp-content/uploads/2022/12/Introducao-a-Estatistica-para-Linguistas.pdf>. Acesso em: 20 abr. 2023.

PAIVA M. C. de. Transcrição de dados linguísticos. In: MOLLICA, M. C.; BRAGA, M. L. (org.). *Introdução à sociolinguística*. São Paulo: Contexto, 2010. p. 117-133.

PAIXÃO DE SOUSA, M. C.; KEPLER, F.; FARIA, P. E-dictor: Novas perspectivas na codificação e edição de corpora de textos históricos. In: VIII Encontro de Linguística de Corpus, 2009. Rio de Janeiro, *Anais do VIII Encontro de Linguística de Corpus*. Rio de Janeiro: UERJ, 2009. p. 69-105.



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

PESSOA, M. de B. *Variação de uma variedade urbana e semi-oralidade: o caso do Recife, Brasil*. Tübingen: Niemeyer, 2003.

PESSOA, Fernando. *Poemas de Alberto Caeiro*. 10. ed. Lisboa: Ática, 1993 [1946]. Disponível em: <http://arquivopessoa.net/typographia/textos/arquivopessoa-1486.pdf>. Acesso em: 19 de ago. de 2021.

PESSOA, Fernando. *Crítica: ensaios, artigos e entrevistas* (edição de Fernando Cabral Martins). Lisboa: Assírio & Alvim, 2000, p. 246-247.

*Perfil dos Territórios de Identidade*. Superintendência de Estudos Econômicos e Sociais da Bahia. 3. v. Salvador: SEI, 2018. Disponível em: <http://www.seplan.ba.gov.br/2018/05/690/Sei-lanca-o-terceiro-volume-do-Perfil-dos-Territorios-de-Identidade-da-Bahia.html>. Acesso em: 14 de fev. de 2021.

PETTER, M. M. T. A classificação das línguas africanas no Brasil. In: PETTER, Margarida M. M. T. (org.). *Introdução à Linguística Africana*. São Paulo: Contexto, 2015a.

PETTER, M. M. T.; ARAÚJO, P. P. Linguística Africana: passado e presente. In: PETTER, M. M. T. (org.). *Introdução à Linguística Africana*. São Paulo: Contexto, 2015b.

PETTER, M. M. T. Aspectos morfosintáticos comuns ao português angolano, brasileiro e moçambicano. *PAPIA* 19, p. 201-220, 2009.

PETTER, M. M. T. Línguas africanas no Brasil África. *Revista do Centro de Estudos Africanos*. USP, São Paulo, 27-28: 63-89, 2006/2007, p. 63-89.

*Plano Territorial de Desenvolvimento Rural Sustentável e Solidário (PTDRSS)*. Conselho de Desenvolvimento Sustentável do Território de Irecê, 2017. Disponível em: [http://sit.mda.gov.br/download/ptdrs/ptdrs\\_qua\\_territorio050.pdf](http://sit.mda.gov.br/download/ptdrs/ptdrs_qua_territorio050.pdf). Acesso em: 15 fev. 2021.

PILATI, E. *Linguística, gramática e aprendizagem ativa*. Campinas: Pontes Editores. 2017.

PILATI, E. N. S.; VENEZIANO, W. H. Sobre o desenvolvimento do aplicativo gramatikê para o ensino de gramática, fundamentos teóricos e questões práticas. *Revista Fórum Linguístico: Ludicidade e Práticas Pedagógicas no Ensino do Português*. v. 20 n. 2, 2023. Disponível em: <https://periodicos.ufsc.br/index.php/forum/article/view/90556>. Acesso em 25 jul. 2023.

*R CORE TEAM*. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria, 2013. URL: <http://www.R-project.org/>.

REIS, João. Quilombos e revoltas escravas no Brasil. *Revista USP*, v. 5, n. 2, São Paulo, 1996, p. 14-39.

RODRIGUES, A. D. As outras línguas da colonização do Brasil. In: CARDOSO, S. A. M.; MOTA, J. A.; MATTOS e SILVA, R. V. (Org.). *Quinhentos anos de história lingüística do Brasil*. Salvador: Secretaria da Cultura e Turismo do Estado da Bahia, 2006. p. 143-162.

RODRIGUES, A, D. Línguas indígenas: 500 anos de descobertas e perdas», *D.E.L.T.A.* 9(1), 83-103, 1993.

RODRIGUES, A. D. *Língua brasileiras: para o conhecimento das línguas indígenas*. São Paulo: Edições Loyola, 1986.

ROSA, J. G. *Grande sertão: veredas*. Rio de Janeiro, José Olympio, 1979.

SANKOFF, D.; TAGLIAMONTE S.; SMITH E. *Goldvarb X: A variable rule application for Macintosh and Windows*. Department of Linguistics, University of Toronto, 2005.

SANTANA, J. H. dos S.; ARAÚJO, S. de F.; FREITAG, R. M. K. Documentação do português falado em comunidades rurais afro-brasileiras de Sergipe: procedimentos metodológicos. *PAPIA*, São Paulo, 28(2), p. 219-237, Jul/Dez 2018a. Disponível em: <http://revistas.fflch.usp.br/papia/article/view/219-237>. Acesso em: 22 jan. 2021.

SANTANA, J. H. S. ; ARAUJO, S. S. F. ; FREITAG, R. M. K. . Documentação do português falado em comunidades rurais afro-brasileiras de Sergipe: patrimônio e memória. *Palimpsesto*, v. 17, p. 121-138, 2018b. Disponível em: <https://www.e-publicacoes.uerj.br/index.php/palimpsesto/article/view/36667>. Acesso em: 20 abr. 2021.

SANTIAGO, H. da S. *et al.* CE-DOHS: um banco de dados sociolinguísticos para a história do português brasileiro. *Labor Histórico*, Rio de Janeiro, 7 (Especial), 2021, p. 311-329.

SANTOS, C. S. dos. 2014. *A Chapada Diamantina: uma experiência arlequiniana de desenvolvimento territorial no Brasil*. Dissertação de mestrado, Curso de Mestrado Multidisciplinar e Profissional em Desenvolvimento e Gestão Social, Universidade Federal da Bahia. 2014. 160 f.

SANTOS, G. *O português afro-brasileiro Helvécia-Ba: análise da variável em coda silábica*. 2012. Tese (Doutorado em Letras e linguística) – Instituto de Letras Universidade Federal da Bahia, Salvador, 2012. Disponível em: <https://repositorio.ufba.br/handle/ri/20082>. Acesso em: 15 abr. 2022.

SANTOS, Milton. O dinheiro e o território. SANTOS, Milton et al. *Território, territórios: ensaios sobre o ordenamento territorial*. 2. ed. Rio de Janeiro: DP&A, 2007. p. 7-13.

SANTOS, Milton. *Por uma outra globalização: do pensamento único à consciência universal*. 10. ed. Rio de Janeiro: Record, 2003. 174 p.

SAPIR, E. *A linguagem: uma introdução ao estudo da fala*. Trad. de Joaquim Mattoso Câmara Jr. Rio de Janeiro: Instituto Nacional do Livro, 1954 [1921].

SAQUET, Marcos Aurélio. *Leituras do conceito de território e de processos espaciais*. São Paulo: Expressão Popular, 2009.

SARTRE, J. P. *O imaginário*. São Paulo: Ática, 1979

SAQUET, Marcos Aurélio. *Leituras do conceito de território e de processos espaciais*. São Paulo: Expressão Popular, 2009.

SAUSSURE, F. *Curso de Linguística Geral*. Trad. Antônio Chelini, José Paulo Paes, Izidoro Blikstein. 27. ed. São Paulo: Cultrix, 2006 [1916].

SCHERRE, M. M. P. Sobre a influência de três variáveis relacionadas na concordância nominal em português. In: SILVA, G. M. de O.; SCHERRE, M. M. P. (Orgs.). *Padrões sociolinguísticos - análise de fenômenos variáveis do português falado na cidade do Rio de Janeiro*. Rio de Janeiro: Editora da UFRJ, 1996. p. 41-62

SCHERRE, M. M. P. Aspectos da concordância de número no português do Brasil. *Revista Internacional de Língua Portuguesa (RILP) - Norma e Variação do Português*. Associação das Universidades de Língua Portuguesa. 12: 37-49. dez. de 1994.

SCHERRE, M. M. P.; NARO, A. J.. Duas dimensões do paralelismo verbal no português popular do Brasil. *DELTA* 9(1):1-14, 1993.

SCHERRE, M. M. P. *Reanálise da concordância nominal em português*. Tese de Doutorado em Linguística. Rio de Janeiro, Faculdade de Letras, UFRJ, 1988.

SCHERRE, M. M. P. *A regra de concordância de número no sintagma nominal em português*. Dissertação de Mestrado, Pontifícia Universidade Católica-Rio, 1978.

SCHERRE, M. M. P.; NARO, A. J. Mudança sem mudança: a concordância de número no português brasileiro. *Scripta*, 9(18), 2006, p.107-129. Disponível em: <http://periodicos.pucminas.br/index.php/scripta/article/view/12597>. Acesso em: 25 maio 2022.

SCHERRE, M. M. P.; NARO, A. J. Sobre a concordância de número no português falado do Brasil. In: RUFFINO, G. (org.) *Dialettologia, geolinguística, sociolinguística*. (Atti del XXI Congresso Internazionale di Linguística e Filologia Romanza) Centro di Studi Filologici e Linguistici Siciliani, Università di Palermo. Tübingen: Max Niemeyer Verlag, 5:509- 523, 1998.

SILVA, Giselle M. de O. Coleta de dados. In: MOLLICA, M. C.; BRAGA, M. L. (org.). *Introdução à sociolinguística*. São Paulo: Contexto, 2010. p. 117-133.

SILVA, J. A.; LACERDA, M. F. O. *Falaê PB*. Disponível em: <https://falaepb.wordpress.com/>. Acesso em: 01 set. 2022.

SILVA NETO, S. da. *Introdução ao estudo da língua portuguesa no Brasil*. Rio de Janeiro: Instituto Nacional do Livro, 1963 [1951].

STEHMANN, J. R.; FARIA, F. S.; BRAGIONI, T.. *50 árvores do museu*. Belo Horizonte: Formato Editora, 2019.

STROUD, C.; PERPÉTUA, G.. (eds.). *Panorama do português oral de Maputo*. Maputo: Instituto Nacional do Desenvolvimento da Educação, 1997.

Superintendência de Estudos Econômicos e Sociais da Bahia (SEI). *Perfil dos Territórios de Identidade*. Salvador: SEI, 2018. 3 v. p.252 (Série territórios de identidade da Bahia, v. 3).

Disponível em:

[https://www.sei.ba.gov.br/index.php?option=com\\_content&view=article&id=2000&Itemid=284#:~:text=O%20Perfil%20dos%20Territ%C3%B3rios%20de,suas%20principais%20potencialidades%20e%20vulnerabilidades](https://www.sei.ba.gov.br/index.php?option=com_content&view=article&id=2000&Itemid=284#:~:text=O%20Perfil%20dos%20Territ%C3%B3rios%20de,suas%20principais%20potencialidades%20e%20vulnerabilidades). Acessado em: 18 mar. 2021.

TARALLO, F. *A pesquisa sociolinguística*. São Paulo: Ática, 2007.

TEYSSIER, P. *História da língua portuguesa*. Tradução Celso Cunha. 3 ed. São Paulo: Martins Fontes, 1997.

VENÂNCIO, R. P. Por uma história social do português no Brasil. In: RAMOS, J. M.; ALKMIM, M. A. (orgs.). *Para a história do português brasileiro: estudos sobre a mudança linguística e história social*. vol. 5. Belo Horizonte: FALE/ UFMG, 2007.

XAVIER, Antonio Carlos; Suzana, CORTEZ Orgs. (2003). *Conversas com linguistas*. São Paulo: Parábola Editorial.

WEEDWOOD, B. *História concisa da Linguística*. Trad. Marcos Bagno. São Paulo: Parábola, 2002.

WEINREICH, U.; LABOV, W.; HERZOG, M. I. *Fundamentos empíricos para uma teoria da mudança linguística*. Tradução Marcos Bagno. São Paulo: Parábola, 2006 [1968].



# ELIHS

Estudos Linguísticos e Históricos do Sertão

