



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

**JOSÉ CARLOS ASSUNÇÃO NOVAES**

**O PREENCHIMENTO DO SUJEITO PRONOMINAL NO PORTUGUÊS  
AFRO-BRASILEIRO NA COMUNIDADE QUILOMBOLA DE  
LAGOINHA**

Salvador

2022

JOSÉ CARLOS ASSUNÇÃO NOVAES

**O PREENCHIMENTO DO SUJEITO PRONOMINAL NO PORTUGUÊS  
AFRO-BRASILEIRO NA COMUNIDADE QUILOMBOLA DE  
LAGOINHA**

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

Orientador: Prof. Dr. Carlos Felipe da Conceição Pinto

Salvador

2022

Ficha catalográfica elaborada pelo Sistema Universitário de Bibliotecas (SIBI/UFBA),  
com os dados fornecidos pelo(a) autor(a).

Novaes, José Carlos Assunção

O preenchimento do sujeito pronominal no português  
afro-brasileiro na comunidade quilombola de Lagoinha  
/ José Carlos Assunção Novaes. -- Salvador, 2022.  
226 f.

Orientador: Carlos Felipe da Conceição Pinto.  
Tese (Doutorado - Língua e Cultura) -- Universidade  
Federal da Bahia, Instituto de Letras, 2022.

1. pronome sujeito. 2. português afro-brasileiro.  
3. português popular do Brasil. 4. linguística. 5.  
sociolinguística. I. Pinto, Carlos Felipe da Conceição.  
II. Título.



Ata da sessão pública do Colegiado do PROGRAMA DE PÓS-GRADUAÇÃO EM LÍNGUA E CULTURA (PPGLINC), realizada em 19/12/2022 para procedimento de defesa da Tese de DOUTORADO EM LÍNGUA E CULTURA no. 25, área de concentração Linguística Histórica, do candidato JOSÉ CARLOS ASSUNÇÃO NOVAES, de matrícula 216122457, intitulada O preenchimento do sujeito no português afrobrasileiro na comunidade quilombola de Lagoinha. Às 14:00 do citado dia, <https://meet.jit.si/CarlosFelipePinto>, foi aberta a sessão pelo presidente da banca examinadora Prof. Dr. CARLOS FELIPE DA CONCEICAO PINTO que apresentou os outros membros da banca: Prof. Dr. DANTE EUSTACHIO LUCCHESI RAMACCIOTTI, Profª. Dra. SONIA MARIA LAZZARINI CYRINO, Profª. Dra. MARIA EUGENIA LAMOGGLIA DUARTE e Prof. Dr. EDUARDO FERREIRA DOS SANTOS. Em seguida foram esclarecidos os procedimentos pelo presidente que passou a palavra ao examinado para apresentação do trabalho de Doutorado. Ao final da apresentação, passou-se à arguição por parte da banca, a qual, em seguida, reuniu-se para a elaboração do parecer. No seu retorno, foi lido o parecer final a respeito do trabalho apresentado pelo candidato, tendo a banca examinadora aprovado o trabalho apresentado, sendo esta aprovação um requisito parcial para a obtenção do grau de Doutor. Em seguida, nada mais havendo a tratar, foi encerrada a sessão pelo presidente da banca, tendo sido, logo a seguir, lavrada a presente ata, abaixo assinada por todos os membros da banca.

**Dra. MARIA EUGENIA LAMOGGLIA DUARTE, UFRJ**  
Examinadora Externa à Instituição

**Dr. EDUARDO FERREIRA DOS SANTOS, UNILAB**  
Examinador Externo à Instituição

**DANTE EUSTACHIO LUCCHESI RAMACCIOTTI, UFF**  
Examinador Interno ao Programa

**Dra. SONIA MARIA LAZZARINI CYRINO, UFBA**  
Examinadora Interna ao Programa

**Dr. CARLOS FELIPE DA CONCEICAO PINTO, UFBA**  
Presidente

**JOSÉ CARLOS ASSUNÇÃO NOVAES**  
Doutorando(a)

## AGRADECIMENTOS

À Deus, por ter me permitido chegar até aqui.

Aos meus pais, Deusdedito Ribeiro de Novaes e Eulina Assunção Novaes (*in memoriam*), por toda a dedicação que tiveram comigo durante todo o tempo em que estiveram no plano material.

Aos meus irmãos, Conceição, Socorro, Fátima, Bráulio e Rodrigo, pelo apoio e carinho de sempre.

Ao meu orientador, Prof. Dr. Carlos Felipe Pinto, por me fazer crescer com seus ensinamentos e suas orientações.

À comunidade quilombola de Lagoinha por ter me recebido em suas casas e ter me permitido a gravação dos inquéritos dessa pesquisa.

À Eliete Fonseca Santos e Romilce Rodrigues Silva Marques pelo apoio.

NOVAES, José Carlos Assunção. *O preenchimento do sujeito pronominal no português afro-brasileiro na comunidade quilombola da Lagoinha*. 2022. 226f. Tese (Doutorado em Língua e Cultura) – Instituto de Letras, Universidade Federal da Bahia, Salvador, 2022.

## RESUMO

Um dos aspectos que diferenciam o português brasileiro do português europeu é a realização do sujeito pronominal (eles compraram a casa ~ compraram a casa). De acordo com a Teoria de Princípios e Parâmetros, proposta por Chomsky (1981), o português europeu é uma língua *pro-drop* típica, pois não realiza o sujeito pronominal em determinados contextos. Já o português brasileiro estaria passando por uma fase de transição de língua *pro-drop* para língua não *pro-drop*. Essa mudança possui uma relação com a significativa redução nos paradigmas flexionais do verbo. Este trabalho busca oferecer subsídios que permitam uma melhor compreensão das variações da língua portuguesa no Brasil, através de um mapeamento de aspectos referentes ao preenchimento do sujeito no português popular afro-brasileiro na comunidade quilombola de Lagoinha, localizada no meio rural do município de Nova Canaã, Bahia, Brasil. Compõe-se, inicialmente de uma análise do parâmetro do sujeito nulo, ou *pro-drop*, no âmbito da teoria de princípios e parâmetros proposta por Chomsky (1981). A partir de um estudo sobre o Português Popular do Brasil, analisaremos a história da língua portuguesa no Brasil, desde o momento em que os portugueses aqui chegaram e a implantaram em território brasileiro, analisando a situação inicial, de contato entre várias línguas. A pesquisa tem sua base empírica em uma amostra de fala vernácula recolhida junto a seis moradores da comunidade quilombola da Lagoinha, no Município de Nova Canaã, na região Sudoeste da Bahia. Esses informantes, com pouco ou nenhuma escolaridade, foram distribuídos pelos dois sexos e por três faixas etárias. A análise da realização do sujeito pronominal nesta amostra linguística fundamentou-se nos princípios teóricos metodológicos da Sociolinguística Variacionista (LABOV, 1972). Os resultados indicaram uma situação de leve desfavorecimento do preenchimento do pronome sujeito (48,4%) em comparação ao uso de sujeito nulo (51,6%). Essa situação de variação permite afirmar que o português brasileiro já não se configura mais como uma língua *pro-drop*. Resultados semelhantes podem ser observados nos trabalhos de Lucchesi (2004, 2009d), Duarte, (1995), Novaes (2007, 2021), dentre outros. Ao analisarmos os resultados da variável faixa etária, em que o processo de mudança é observado em tempo aparente, percebemos que os mais jovens (faixa etária II e I) são os que mais realizam o sujeito pronominal pleno.

**Palavras-chave:** Pronome sujeito; Teoria da Gramática; Sociolinguística; Português popular do Brasil; Português afro-brasileiro.

NOVAES, José Carlos Assunção. *Completing the subject pronoun in Afro-Brazilian Portuguese in the quilombola community of Lagoinha*. 2022. 226f. Tese (Doutorado em Língua e Cultura) – Instituto de Letras, Universidade Federal da Bahia, Salvador, 2022.

## ABSTRACT

One of the aspects that differentiate Brazilian Portuguese from European Portuguese is the realization of the pronominal subject. According to the Theory of Principles and Parameters, proposed by Chomsky (1981), European Portuguese is a typical *pro-drop* language. This definition is given when the pronominal subject is not realized in certain contexts. Brazilian Portuguese, on the other hand, would be going through a phase of transition from a *pro-drop* language to a *non-pro-drop* language. This change is related to the significant reduction in the inflectional paradigms of the verb. This work seeks to provide a better understanding of the variations of the Portuguese language in Brazil, through a mapping of aspects related to the completion of the subject in popular Afro-Brazilian Portuguese in the quilombola community of Lagoinha, located in the rural area of the municipality of Nova Canaã, Bahia. Initially, it consists of an analysis of the null subject parameter (or *pro-drop* parameter) within the scope of the Theory of Principles and Parameters, proposed by Chomsky (1981), in Universal Grammar. Based on a study on Popular Portuguese in Brazil, we analyze the history of the Portuguese language in Brazil, since the moment that Portuguese arrived in the country and was implanted in Brazilian territory, investigating the initial situation of contact between several languages. The research has its empirical basis in a sample of vernacular speech collected from six residents of the quilombola community of Lagoinha, in the Municipality of Nova Canaã, in the Southwest region of Bahia. These subjects, all with low level of education, were distributed by both sexes and by three age groups. The analysis of the realization of the pronominal subject in this linguistic sample was based on the Variationist Sociolinguistics (LABOV, 1972). The results indicated a situation of slight disadvantage in filling out the subject pronoun (48,4%) in comparison to the null subject variant (51,6%). This situation of variation in Brazilian Portuguese shows that it cannot be classified as a *pro-drop* language. Similar results can also be observed in Lucchesi (2004, 2009d), Duarte (1995) and Novaes (2007, 2021). When analyzing the results of the age group variable, in which the change process is observed in apparent time, we noticed that the youngest (age group II and I) are the ones who most perform the full pronominal subject.

**Keywords:** Subject Pronoun; Grammar Theory; Sociolinguistics; Popular Brazilian Portuguese; Afro-Brazilian Portuguese.

## LISTA DE FIGURAS

<b>Figura 1</b> - Princípio do subconjunto .....	78
<b>Figura 2</b> - Moradores da Comunidade da Lagoinha, após reunião da Associação.	165
<b>Figura 3</b> - José Carlos Assunção Novaes com Eliete Fonseca Santos e Lídio Fonseca Araújo, moradores da Comunidade da Lagoinha .....	166
<b>Figura 4</b> - Reunião da Associação.....	166
<b>Figura 5</b> - Moradora mostra forma artesanal de produção de farinha na comunidade .....	166
<b>Figura 6</b> - Sede da Associação de Moradores da Lagoinha.....	166
<b>Figura 7</b> - José Carlos Assunção Novaes e moradores da comunidade .....	167
<b>Figura 8</b> - Índices percentuais de realização de pronome sujeito preenchido e de sujeito nulo .....	182
<b>Figura 9</b> - Índices percentuais referentes à primeira menção do sujeito, segundo a pessoa do discurso .....	186
<b>Figura 10</b> - Índices percentuais do uso do pronome de 2ª pessoa do singular ....	1921
<b>Figura 11</b> - Índices percentuais do uso do pronome de 1ª pessoa do plural .....	1922
<b>Figura 12</b> - Índices percentuais de realização de pronome sujeito preenchidos e de sujeito nulo de acordo com o sexo. ....	2088

## LISTA DE TABELAS

<b>Tabela 1</b> - A realização do sujeito pronominal no português afro-brasileiro na comunidade quilombola de Lagoinha – Nova Canaã – BA .....	181
<b>Tabela 2</b> - A realização do sujeito pronominal expresso no português afro-brasileiro na comunidade quilombola de Lagoinha – Nova Canaã – BA, segundo a variável menção anterior do sujeito. ....	185
<b>Tabela 3</b> - Variável desinência gramatical no português afro-brasileiro na comunidade quilombola de Lagoinha – Nova Canaã – BA .....	190
<b>Tabela 4</b> - Variável traço semântico do sujeito no português afro-brasileiro na comunidade quilombola de Lagoinha – Nova Canaã – BA .....	195
<b>Tabela 5</b> - Variável sexo no português afro-brasileiro na comunidade quilombola de Lagoinha – Nova Canaã – BA.....	1988
<b>Tabela 6</b> - Variável escolaridade no português afro-brasileiro na comunidade quilombola de Lagoinha – Nova Canaã – BA.....	201
<b>Tabela 7</b> - A realização do sujeito pronominal expresso no português afro-brasileiro na comunidade quilombola de Lagoinha Nova Canaã – BA, segundo o cruzamento pessoa do discurso vs. desinência verbal .....	2077
<b>Tabela 8</b> - A realização do sujeito pronominal expresso no português afro-brasileiro na comunidade quilombola de Lagoinha – Nova Canaã – BA, segundo a variável sexo.....	210

# SUMÁRIO

<b>INTRODUÇÃO .....</b>	<b>11</b>
<b>1 RELAÇÕES ENTRE A teoria da gramática e a sociolinguística .....</b>	<b>18</b>
1.1 TEORIA DA GRAMÁTICA .....	18
1.1.1 Língua-E e língua-I .....	22
1.2 A SOCIOLINGUÍSTICA.....	30
1.2.1 Labov e a Sociolinguística .....	35
1.2.2 Sociolinguística Paramétrica.....	41
1.3 CONCLUINDO A SEÇÃO .....	49
<b>2 TEORIA DE PRINCÍPIOS E PARÂMETROS .....</b>	<b>51</b>
2.1 A MUDANÇA DE UM SISTEMA DE REGRAS PARA UM MODELO FUNDAMENTADO EM PRINCÍPIOS E PARÂMETROS.....	51
2.2 PARÂMETROS E GATILHOS NO PROCESSO DE AQUISIÇÃO DA LINGUAGEM.....	58
2.3 O PARÂMETRO DO SUJEITO NULO .....	68
2.3.1 O Princípio do Subconjunto e o acesso a dados negativos durante o processo da aprendizagem.....	76
2.4 O SUJEITO NULO NO PORTUGUÊS DO BRASIL.....	81
2.4.1 O caráter variável da concordância sujeito-verbo e o parâmetro do sujeito nulo no PB .....	94
2.5 CONCLUINDO A SEÇÃO .....	102
<b>3 O PORTUGUÊS AFRO-BRASILEIRO .....</b>	<b>104</b>
3.1 OS AFRICANOS NA FORMAÇÃO DA SOCIEDADE BRASILEIRA E O PLURILINGUISMO AFRICANO NO BRASIL .....	113
3.2 A LÍNGUA PORTUGUESA FALADA PELOS AFRICANOS ESCRAVIZADOS NO BRASIL E A INFLUÊNCIA DAS LÍNGUAS AFRICANAS NA FORMAÇÃO DO PORTUGUÊS DO BRASIL .....	130
3.3 CRIOLIZAÇÃO, SEMI-CRIOLIZAÇÃO E TRANSMISSÃO LINGUÍSTICA IRREGULAR .....	137
3.3.1 Transmissão linguística irregular, nativização e a polarização sociolinguística do Brasil .....	148
3.4 CONCLUINDO A SEÇÃO .....	160

<b>4 MÉTODOS</b> .....	<b>162</b>
4.1 DESCRIÇÃO DA COMUNIDADE ESTUDADA .....	162
4.2 O ESTUDO DO USO DA LÍNGUA DENTRO DA COMUNIDADE DE FALA.....	167
4.3 METODOLOGIA .....	171
4.3.1 Amostras de falas: entrevistas .....	171
4.3.2 Caracterização dos informantes.....	173
4.3.4 Gravação dos inquéritos de fala.....	175
4.3.5 Processamento quantitativo dos dados.....	177
4.4 CONCLUINDO A SEÇÃO .....	178
<b>5 ANÁLISE DE DADOS</b> .....	<b>179</b>
5.1 A VARIÁVEL DEPENDENTE.....	181
5.2 VARIÁVEIS INDEPENDENTES LINGUÍSTICAS .....	183
5.2.1 Menção anterior do sujeito.....	184
5.2.2 Pessoa do discurso.....	188
5.2.3 Traço semântico do sujeito .....	194
5.2.4 Desinência gramatical do verbo .....	197
5.2.5 Cruzamento da pessoa do discurso com a desinência gramatical do verbo.....	2000
5.2.6 Conclusão sobre as variáveis linguísticas .....	205
5.3 VARIÁVEIS INDEPENDENTES EXTRALINGUÍSTICAS.....	206
5.3.1 Sexo .....	207
5.3.2 Faixa etária .....	210
5.3.3 Conclusão das variáveis sociais .....	211
5.4 CONCLUSÃO DA ANÁLISE DE DADOS .....	2111
<b>CONCLUSÕES</b> .....	<b>214</b>
<b>REFERÊNCIAS</b> .....	<b>218</b>

## INTRODUÇÃO

A língua portuguesa foi transplantada para o Brasil a partir do século XVI, encontrando, em território brasileiro, várias outras línguas sendo utilizadas pelos indígenas. Além disso, com a vinda de mão-de-obra escrava para o país, também se deu o contato entre falantes do português com os de línguas africanas diversas. O português falado no Brasil não sofreu as mudanças ocorridas a partir do século XVIII na língua falada na metrópole e manteve-se mais próximo do português falado em Portugal na época da descoberta, com fortes influências indígenas, africanas e, mais tarde, de imigrantes europeus que se instalaram no Centro-Sul.

A vinda dos escravizados, oriundos de diversas regiões da África, no início de nossa colonização, é um dos fatores relacionados à nossa história linguística. Discute-se muito que influência essa população teria em relação à formação do nosso português popular, falado pelas camadas da população não escolarizada e socialmente desligada dos padrões cultos da nossa língua, principalmente aquelas remanescentes de quilombos ou situadas na zona rural, longe dos centros urbanos e, portanto, isoladas culturalmente. No português popular do Brasil, os processos de transmissão linguística irregulares mais leves compreenderam uma fase inicial, segundo Lucchesi (2003), na qual ocorreu uma variação mais ou menos intensa em relação ao uso dos mecanismos gramaticais e, nas fases seguintes, uma manutenção desse quadro de variação, o que resultou numa heterogeneidade da fala.

Além de usarem o português como uma segunda língua, modificada pelas condições de aquisição, os escravizados e ex escravizados foram transmitindo essa língua, um português modificado, para os seus descendentes, que a receberam como sua primeira língua. É claro que, sendo maioria em algumas comunidades, podem ter influenciado em alguns traços do português popular, principalmente nos da zona rural, local para onde eram enviados para o trabalho nas plantações de cana-de-açúcar ou de fumo. Essa população, além disso, não tinha tanto contato com as vilas e povoados que existiam na época.

A língua oficial do Brasil é a língua portuguesa, cujo padrão é tomado pela escola como modelo para o ensino. No entanto, constata-se a utilização de um português variável, por vezes, bem diferente da orientação escolar, o que é observado principalmente em comunidades rurais no interior do Brasil. O português

brasileiro (PB), sobretudo em suas variedades populares, diferencia-se muito do português europeu (PE), pois possui características próprias que se formaram desde o período da colonização do Brasil, quando aqui chegaram os primeiros colonizadores europeus e estabeleceram contato com as diversas línguas indígenas e, posteriormente, com as línguas africanas dos escravizados trazidos do continente africano. Esse contato teria afetado principalmente a formação do português popular do Brasil, não deixando de se refletir também na norma culta brasileira.

Desse modo, alguns aspectos da variação do português rural são comuns a diversas regiões, outros, conforme Baxter (1992) são particularidades de comunidades negras isoladas, resultantes de transmissão irregular do português. Dentre esses aspectos, o autor inclui a redução da morfologia verbal de pessoa, tipo de ocorrência pouco documentada em comunidades quilombolas no interior do estado da Bahia.

Nesse sentido, no Brasil, ainda há comunidades rurais, conforme Lucchesi (2012), originadas de antigos agrupamentos de escravizados foragidos, originados de antigos quilombos. Essas comunidades são sítios arqueológicos da história sociolinguística do Brasil e, por isso mesmo, o estudo da fala de seus moradores é de grande importância, quando se procura explicar as variedades populares do PB, observando-se o contato entre línguas durante a sua formação histórica, já que elas seriam aquelas mais afetadas pelo processo de transmissão linguística irregular.

Sendo assim, é possível encontrar certas características próprias desse contato como, por exemplo, as que se referem ao parâmetro do sujeito nulo no português popular brasileiro. Ponte (2008) relata, por exemplo, o caso das línguas crioulas, em que a ausência de morfologia de pessoa e número torna obrigatório o uso de pronomes realizados, ocorrendo até em crioulos que têm como base as línguas românicas.

Em relação às mudanças que ocorrem na língua, abordaremos também um dos fenômenos mais discutidos atualmente, que diz respeito à Teoria de Princípios e Parâmetros, proposta por Chomsky (1981), quando se iniciou nos estudos da Gramática Gerativa, uma fase marcada pela busca dos princípios universais das línguas humanas.

A espécie humana possui capacidades que são inatas, sendo a faculdade da linguagem também uma das capacidades que fazem parte da herança genética do ser humano, assim como a visão, a audição etc. Ela, assim como as outras

faculdades humanas, possui uma estrutura caracterizada por propriedades determinadas geneticamente, possibilitando a aprendizagem, que tem como ponto de partida o estado mental inicial da linguagem. Esse estado inicial passaria por vários estágios de desenvolvimento, sofrendo a influência do meio e das experiências pessoais, através de um processo de maturação que atingiria um estágio de estabilidade, o que não impede que modificações posteriores venham a ocorrer, apesar de serem, nesse caso, diferentes das que se processam entre a aquisição, no estado inicial e a estabilidade. Esse estado mental inicial da linguagem é chamado por Chomsky (1988 [1981]) de *Gramática Universal* (GU).

Aos princípios universais da linguagem humana, estariam associados parâmetros variáveis, ou seja, propriedades sintáticas que admitiriam duas possibilidades de manifestação. Dentro desse contexto, temos o parâmetro do sujeito nulo ou *pro-drop*. Quando uma língua é marcada positivamente em relação a esse parâmetro, pode admitir tanto sujeitos referenciais quanto sujeitos não referenciais, ou seja, foneticamente nulos. O português europeu, assim como o italiano e o espanhol, é uma língua *pro-drop*, permitindo as duas possibilidades de pronome sujeito. No entanto, o português brasileiro, devido à perda da morfologia flexional, está se tornando uma língua *não-pro-drop* (DUARTE, 1995; XAVIER, 2006; KATO, 2000).

Desde a década de 1960, um debate existente na Linguística Moderna embasa a ideia de que os modelos teóricos da Sociolinguística Variacionista (SV) e da Gramática Gerativa (GG) não são inflexíveis entre si. Isso, segundo, é baseado na utilização de conceitos comuns, que são cruciais para se construir o objeto de estudo em cada um dos dois paradigmas. O autor também afirma que isso fica evidente em um dos principais trabalhos da SV, os Fundamentos Empíricos para uma teoria da Mudança Linguística, de Weinreich, Labov e Herzog (2006), que foi publicado originalmente em 1968. Esses autores enfatizam que a mudança linguística é concebida racionalmente através da descrição ordenada da diferenciação em uma língua que serve a uma comunidade. Sendo assim, eles defendem que: “o comando nativo das estruturas heterogêneas não é matéria de multidialectalismo ou de ‘mero’ desempenho, mas é parte da competência linguística monolíngue” (WEINREICH; LABOV; HERZOG, 2006, p. 36).

O fato de os sociolinguistas terem adotado conceitos como competência e desempenho, considerados centrais na reflexão dos gerativistas, e o termo *code-*

*switching*, para indicar a alternância de código, ao discutir a natureza do sistema linguístico, é uma indicação de que o diálogo entre os dois modelos é uma realidade objetiva. Os modelos teóricos da SV e da GG estariam reunidos decisivamente num mesmo campo da ciência da linguagem, segundo Lucchesi (2011, p. 232), através da “adoção da noção de estrutura (ou sistema) como um elemento chave em qualquer algoritmo explicativo da linguagem humana”. Ao reunir os dois modelos, estes estariam opostos, então, à Análise da Conversação e à Análise do Discurso, correntes das linguísticas contemporâneas definidas como pós-modernas.

A partir desse antagonismo existente no cenário da linguística contemporânea, considerados irreduzíveis entre si, entre a Linguística Moderna e a Linguística Pós-Moderna, a GG e a SV estariam juntas no campo da modernidade:

Colocamos a Gramática Gerativa e Sociolinguística Variacionista no campo da modernidade, como as duas vertentes que sucedem o estruturalismo linguístico, negando-o, mas mantendo-se no campo de uma ciência da linguagem que pretende uma apreensão global e sistemática do fenômeno linguístico através do conceito de uma estrutura (ou sistema) que possui uma lógica própria (LUCCHESI, 2011, p. 233).

Esses conceitos foram aplicados também em pesquisa prévia (NOVAES, 2007, 2021) sobre o Parâmetro do sujeito nulo no Português Popular do interior do estado da Bahia, no município de Poções, situado na região Sudoeste da Bahia, comparando a realização das falas na sede do município (zona urbana) e no distrito de Morrinhos (zona rural). Os resultados da análise variacionista do sujeito pronominal nulo, de referência definida no português popular do interior do estado da Bahia, revelaram uma semelhança entre os condicionantes linguísticos que regem a variação na fala de comunidades afro-brasileira (LUCCHESI, 2004, 2009d) e na variedade urbana culta (DUARTE, 1995) do PB. Tanto nos estudos de Duarte (1995) e Lucchesi (2004, 2009d), como em Novaes (2007, 2021), o sujeito nulo é favorecido pela presença de um morfema flexional do verbo, pelo mecanismo sintático da correferência e pelo traço semântico [-humano]. Lucchesi (2004, 2009d) afirma que tais resultados fortalecem a visão de que existe uma universalidade nas ações e nos processos de variação e mudança no sistema linguístico. Isso ocorre tanto no plano diacrônico quanto sincrônico, bem como em suas variações diastráticas.

Observamos, no entanto, algumas peculiaridades no plano sociolinguístico entre esses estudos. Percebemos que há uma similaridade em relação ao uso das variantes na frequência do uso do sujeito pronominal nulo na norma culta, no português afro-brasileiro e no português popular. No entanto, as tendências de mudança observadas são, segundo os dados analisados, diferentes. Tanto na norma culta (DUARTE, 1995) quanto no português popular (NOVAES, 2007, 2021), foi observada uma tendência para a perda da propriedade do sujeito pronominal nulo. Já na análise de Lucchesi (2004, 2009d) sobre o português afro-brasileiro, não se chegou a um resultado conclusivo. Em sua análise, ele afirma que estaria em curso, na norma culta, uma queda na frequência de utilização do sujeito nulo, que foi iniciada através de um processo de reorganização na pauta dos pronomes sujeito. Já nas variedades vernáculas, havia uma restrição, em graus diferenciados, de acordo com o nível de erosão de sua morfologia flexional. Tal fato foi determinado pelo grau de contato entre línguas na história dessas variedades.

O PB, tanto em sua norma culta (DUARTE, 1995) e popular (NOVAES, 2007, 2021), e em nossos estudos, no português popular afro-brasileiro, estaria se afastando do grupo de línguas consideradas *pro-drop*? Apesar de ainda conviver com um sistema que reflete as características *pro-drop*, o PB apresenta um sistema em desenvolvimento em que a perda da “riqueza funcional” já não permite mais a identificação como uma língua com características *pro*?

Para responder a essas perguntas, surgiu o interesse de realizarmos essa pesquisa, observando a questão do preenchimento do pronome sujeito no português popular afro-brasileiro na comunidade quilombola de Lagoinha, o que permitiria também uma comparação com o resultado obtido anteriormente nos trabalhos de Novaes (2007, 2021) sobre o português popular do interior do estado da Bahia, devido à proximidade das comunidades que foram e serão objetos de estudo. Estaremos também comparando os resultados obtidos com os trabalhos de Duarte (1995) e Lucchesi (2004, 2009d).

O objetivo do nosso trabalho é, então, mapear aspectos referentes ao parâmetro do sujeito nulo no português popular afro-brasileiro, na tentativa de identificar motivações linguísticas e não linguísticas para o uso variável da morfologia verbal de pessoa e verificar o uso de pronomes realizados, de acordo com a teoria do parâmetro do sujeito nulo, na variedade de Lagoinha, bem como relacionar a variação linguística existente em Lagoinha e a sócio-história dessa

comunidade e, além desses elementos, oferecer subsídios que permitam uma maior compreensão das variações da língua portuguesa no Brasil para um melhor posicionamento sobre a formação de um português brasileiro. A relevância desta pesquisa justifica-se em função de ainda haver muitas questões a serem discutidas quando se trata do parâmetro do sujeito nulo no português afro-brasileiro. A nossa hipótese inicial é de que encontraremos um maior preenchimento do pronome sujeito nos dados obtidos, ao analisarmos inicialmente os resultados dos trabalhos citados acima e por se tratar de uma comunidade afrodescendente remanescente de quilombo.

O nosso trabalho de pesquisa se compõe, inicialmente, de um estudo sobre o Parâmetro do Sujeito Nulo, no âmbito da Teoria de Princípios e Parâmetros proposta por Chomsky (1981). Assim, abordaremos o parâmetro do sujeito nulo na GU e o parâmetro *pro-drop*, que ocorrem nas línguas em que é dispensável a realização fonética do sujeito pronominal. Essas línguas são chamadas *pro-drop*, ou línguas com *sujeito nulo*, por permitirem a opção da realização fonética ou não do pronome sujeito.

As teorias a respeito do estudo da variação e mudanças linguísticas também serão abordadas neste trabalho de pesquisa. Mostraremos como começaram os primeiros estudos relacionados a esse tema, abordando ainda o conceito de norma linguística. A partir daí, verificaremos como se iniciaram os estudos sobre a variação linguística, a relação entre língua e sociedade e a contribuição da Sociolinguística para os trabalhos mais recentes sobre a língua.

A partir daí, falaremos sobre o Português Popular do Brasil, através de uma análise da história da língua portuguesa no Brasil desde o momento em que os portugueses aqui chegaram e a implantaram em território brasileiro. Analisaremos a situação inicial em que o PE entrou em contato com as várias línguas faladas pelos indígenas que habitavam o Brasil, até a chegada dos negros africanos de diversas regiões da África, que também falavam línguas diversificadas, de acordo com o seu local de origem, de onde foram trazidos como escravos. Abordaremos como a língua portuguesa se estabeleceu em terras brasileiras e as principais variações e diferenças entre o Português do Brasil e o Português de Portugal.

Logo a seguir, faremos um breve relato histórico da comunidade estudada, descrevendo as amostras de fala e os tipos de entrevista, além do processamento de dados realizado para se obter os resultados que serviram de base para a nossa

análise. Também abordaremos a metodologia empregada para o recolhimento de dados para a constituição do *corpus* a ser analisado.

Mostraremos os resultados obtidos através de nossa análise de dados da variável dependente, das variáveis linguísticas e das variáveis sociais, observando os fatores que foram mais relevantes e que demonstraram uma maior significância para a conclusão de nossos estudos. Teremos como variável dependente a realização ou não do sujeito pronominal em que serão analisadas as variáveis linguísticas: pessoa do sujeito, desinência gramatical do verbo e caracterização semântica do sujeito. Quanto às variáveis sociais, analisaremos aquelas referentes a sexo, faixa etária e escolaridade, não levando em consideração o local de moradia, já que todos os informantes são moradores da zona rural, na comunidade quilombola de Lagoinha.

Finalizando o nosso trabalho de pesquisa, apresentaremos uma síntese dos resultados obtidos em nossos estudos e a nossa interpretação final sobre o preenchimento do pronome sujeito no português popular afro-brasileiro na comunidade quilombola de Lagoinha.

## 1 RELAÇÕES ENTRE A TEORIA DA GRAMÁTICA E A SOCIOLINGUÍSTICA

Nesta seção, serão abordados aspectos referentes às relações entre a Teoria da Gramática e a Sociolinguística. Para tanto, será feito um breve relato do surgimento da teoria gerativa de Chomsky até a sociolinguística de Labov, buscando mostrar a evolução dos estudos linguísticos, a partir da visão desses autores.

Na subseção 1.1, abordaremos os estudos relativos à Teoria da Gramática, onde discorreremos sobre a Gramática Universal e o conceito de Língua-E e Língua-I. Na subseção 1.2, por sua vez, voltaremos nossa atenção para a sociolinguística, mais especificamente a de orientação variacionista, assim como para o seu principal difusor, William Labov e, posteriormente, sobre a sociolinguística paramétrica.

### 1.1 TEORIA DA GRAMÁTICA

Quando se fala nos estudos referentes à Gramática Gerativa, logo se associa a esta o nome do linguista norte-americano Noam Chomsky, já que é tema central nos estudos e pesquisas gerativistas, sendo o autor um precursor, desde a publicação, em 1957, de *Syntactic Structures*, em que assumiu o aspecto biológico da linguagem como objeto principal de sua pesquisa. Como em todo procedimento científico, os estudos gramaticais à luz da teoria gerativa surgiram do rompimento com outras teorias, principalmente com o estruturalismo americano e europeu. Os estudos linguísticos, a partir de Chomsky (2015 [1957]), passam a ser vistos, então, através de uma nova abordagem, sendo que não têm como foco a descrição do sistema em si, mas um estudo de base mentalista, girando em torno da faculdade da linguagem. Ao contrário do exposto por Saussure (2006 [1916]), a teoria gerativa afirma que a capacidade para a linguagem é inata ao ser humano, preocupando-se, desse modo, com a explicação e funcionamento de seu aspecto biológico, assim como com a competência linguística do falante.

Os trabalhos gerativistas podem ser divididos em três períodos, de acordo com Boeckx e Hornstein (2007). Todas essas fases têm uma relação com trabalhos de Chomsky e com as pesquisas realizadas em cada período, o que os distingue em relação aos aspectos técnicos e teóricos. Sendo assim, o primeiro desses possui uma relação mais direta com a engenharia e se refere à fase combinatória. A essa fase, estaria associado o *Syntactic Structures* (CHOMSKY, 1957). Já o segundo

período, a fase cognitiva, possui uma relação com a biologia, estando associada a *Aspects of the Theory of Syntax* (CHOMSKY, 1965) e *Lectures on Government and Binding* (CHOMSKY, 1981). Por último, o terceiro período, o minimalista, estaria relacionado com a física, tendo relação com *The Minimalist Program* (CHOMSKY, 1995). Com mais de seis décadas de pesquisas linguísticas, os estudos gerativistas, como é normal acontecer em qualquer programa científico, passaram por mudanças em relação a seus objetivos e métodos de programa.

O estudo do conhecimento que é disponível a qualquer falante nativo, ou seja, o estudo da “língua” abstrata recebeu um novo impulso, a partir de Chomsky (1978 [1965]), ao opor a *competência* ao *desempenho*. Enquanto a primeira se refere ao conhecimento abstrato das regras da língua, o segundo diz respeito à seleção e execução dessas regras. A Teoria Gerativa, para o autor, se refere ao estudo da competência, o que deixa explícita a prática decorrente do paradoxo saussuriano em que o objeto real do estudo linguístico é uma comunidade abstrata, homogênea, na qual todos falam iguais e aprendem a língua de forma instantânea, já que:

A teoria linguística tem antes de mais nada como objecto um falante-ouvinte ideal, situado numa comunidade linguística completamente homogênea, que conhece a sua língua perfeitamente, e que, ao aplicar o seu conhecimento da língua numa performance efectiva, não é afectado por condições gramaticalmente irrelevantes tais como limitações de memória, distrações, desvios de atenção e de interesse, e erros (casuais ou característicos) (CHOMSKY, 1978 [1965], p. 84).

Outro aspecto a se observar na linguística, não deve ser o enunciado do indivíduo a ser estudado, conforme Chomsky (1978 [1965]), mas suas percepções sobre a língua, no que diz respeito, principalmente, aos seus julgamentos acerca de quais frases são gramaticais e quais não são, além da relação entre essas frases.

A gramática gerativa de uma língua particular, segundo Chomsky (1994, p. 22): “é uma teoria que diz respeito à forma e ao significado de expressões dessa língua”. Há muitos modos de se abordar tais questões e, além disso, essa discussão pode ser abordada a partir de diferentes pontos de vista. Limitando-se a considerar certos elementos dessa problemática, a gramática gerativa tem como ponto de referência a psicologia individual e está relacionada aos aspectos da forma e do significado determinados por um componente particular da mente humana denominada de faculdade da linguagem. A essência dessa faculdade é o principal tema de uma teoria geral, chamada de Gramática Universal (GU), sobre a estrutura

da linguagem, cujo objetivo é a descoberta do conjunto de princípios e de elementos comuns às línguas possíveis:

A GU deve ser vista como uma caracterização da faculdade da linguagem geneticamente determinada. Pode-se encarar esta faculdade como um “mecanismo de aquisição da linguagem”, uma componente inata da mente humana que origina uma língua particular pela interação com a experiência vivida, ou ainda como um mecanismo que converte a experiência num sistema de conhecimento atingido: o conhecimento de uma ou de outra língua (CHOMSKY, 1994 [1986], p. 23).

A abordagem da linguagem passou por uma mudança significativa, adquirindo uma nova perspectiva a partir do estudo da gramática gerativa. Isso porque houve o direcionamento do objeto de investigação do comportamento linguístico ou dos produtos desse comportamento para os estados da mente/cérebro que produzem tal comportamento. Sendo assim, o princípio fundamental passa a ser o conhecimento da língua: a sua natureza, a sua origem e o seu uso. Dentro desse contexto, são colocadas três questões básicas que dizem respeito ao que constitui, bem como a como é adquirido e a como é usado o conhecimento da língua.

A resposta para a primeira questão, ou seja, o que constitui o conhecimento da língua está em uma gramática gerativa particular: “uma teoria que se ocupa do estado da mente/cérebro do indivíduo que conhece uma língua particular” (CHOMSKY, 1994 [1986], p. 23). Quanto ao modo como é adquirido o conhecimento da língua, a resposta estaria em uma especificação da GU e “pela consideração dos meios através dos quais os seus princípios interagem com a experiência, de modo a darem origem a uma língua particular” (CHOMSKY, 1994 [1986], p. 23). A GU seria então “uma teoria do “estado inicial” da faculdade da linguagem, anterior a qualquer experiência linguística” (CHOMSKY, 1994 [1986], p. 23). Já sobre a terceira questão, que diz respeito a como é usado o conhecimento da língua, a resposta estaria em “uma teoria acerca do modo como o conhecimento da língua atingido interfere na expressão do pensamento e na compreensão das amostras de línguas que nos são apresentadas e, por consequência, na comunicação e noutros usos especiais da língua” (CHOMSKY, 1994 [1986], p. 23).

Ao admitirmos que há uma faculdade da linguagem, considerada um componente da mente/cérebro que origina o conhecimento da língua e a experiência vivida, conforme afirma Chomsky (1994 [1986]), devemos levar em conta também as questões abordadas acima. Desse modo, busca-se estabelecer qual é o sistema de

conhecimento a ser atingido e quais seriam as propriedades atribuídas ao estado inicial da mente/cérebro para perceber como é alcançado esse conhecimento. Sendo assim, há uma faculdade da linguagem distinta, a partir do momento em que essas propriedades são específicas da linguagem, tanto individualmente como pelo modo como se organizam.

As respostas a essas questões têm avançado, desde meados dos anos de 1950, de acordo com Chomsky (1994 [1986]), o que motivou a criação de um programa de investigação com o intuito de analisar a adequação dessas propostas para que pudessem ser desenvolvidas e aplicadas. O desenvolvimento das ciências cognitivas, na forma como é vista na atualidade, foi conduzido por esse programa, que dividiu com outras abordagens a convicção de que determinados aspectos da mente/cérebro podem ser estudados de maneira proveitosa “tendo como modelo sistemas computacionais de regras que formam e modificam representações e que são usados na interpretação e na acção” (CHOMSKY, 1994 [1986], p. 25).

Desse modo, a GU pode ser concebida como um órgão biológico, inato ao indivíduo e que amadurece como qualquer outro órgão. Essa evolução vai levar à gramática final, caracterizando todos os conhecimentos linguísticos do falante adulto. A linguística, para a gramática gerativa, é considerada, então, uma ciência natural, conforme Xavier (2006, p. 85), cujo objeto de estudo é a faculdade da linguagem. Essa, por sua vez, é um componente da mente/cérebro específico para a linguagem humana. A Gramática Universal é o estado inicial da faculdade da linguagem ( $S_0$ ), conforme Chomsky (1978 [1965]), e a gramática do indivíduo adulto constitui o seu estado final, firme ou estável ( $S_s$ , do inglês *Steady State*).

Assim, estariam contidos, então, na Gramática Universal, princípios linguísticos determinados geneticamente que seriam específicos à espécie humana que, segundo Chomsky (1978 [1965]) permitiria que um falante aprendiz de L1, com base na experiência linguística a que está exposto, chegasse à gramática da sua língua. Sendo assim, as propriedades da gramática final serão explicadas através das propriedades da Gramática Universal, que são extremamente importantes, em qualquer fenômeno gramatical, para determiná-las. Desse modo, o Gerativismo tem como objetivo principal estudar a natureza e as propriedades da Gramática Universal da forma mais precisa possível.

Chomsky (1978 [1965]) defende, então, uma teoria linguística que contenha uma definição de gramática como uma especificação da classe de gramáticas

potenciais: “Correlativamente, podemos dizer que *uma teoria linguística é descritivamente adequada* se fornecer uma gramática descritivamente adequada para cada língua natural” (CHOMSKY, 1978 [1965], p. 107). Sendo assim:

Ainda que até a adequação descritiva em larga escala não seja de modo nenhum fácil de atingir, torna-se crucial para o desenvolvimento produtivo da teoria linguística tentar alcançar objectivos muito mais elevados do que este. Para facilitar a formulação clara de questões mais profundas, é útil considerar o problema abstracto de construir um “modelo de aquisição” para a linguagem, isto é, uma teoria da aprendizagem da linguagem ou da construção de gramáticas. Manifestamente, uma criança que aprendeu uma língua desenvolveu uma representação interna de um sistema de regras que determinam como é que as frases se formam, se utilizam, e se compreendem (CHOMSKY, 1978 [1965], p. 107).

O termo “gramática” é utilizado por Chomsky (1978 [1965]), então, referindo-se, primeiramente, à teoria que o falante nativo possui de sua língua, que é representada internamente, e à aplicação que dela faz o linguista, para afirmar que a criança que aprendeu uma língua “desenvolveu e representou internamente uma gramática generativa, no sentido descrito” (CHOMSKY, 1978 [1965], p.107).

Para um melhor entendimento dos pontos abordados anteriormente, faz-se necessária à exposição das noções de Língua-E e Língua-I, as quais permitirão um melhor esclarecimento sobre os estudos de Chomsky e da Teoria da Gramática.

### 1.1.1 Língua-E e língua-I

Segundo Chomsky (1994 [1986]), várias noções técnicas de “língua” têm sido desenvolvidas através de algumas abordagens científicas, para substituir a noção do senso comum desse termo, que se baseia em uma dimensão sociopolítica que “inclui os conceitos de conhecimento ou compreensão da língua, regra da língua, comportamento linguístico orientado pelas regras e outros” (CHOMSKY, 1994 [1986], p. 35), que está fundamentada na ideia de uma língua homogênea, pura e idealizada, em que prevalece às noções de fala “correta”, de “erro” e de fidelidade à “gramática”. Essa situação é refletida na sociedade que vê a língua como um sistema homogêneo e unitário:

Tal situação reflete a visão hegemônica na sociedade, segundo a qual a língua é um sistema homogêneo e unitário, desenvolvido pelos melhores escritores nacionais e fixado pelos gramáticos, para ser seguido e usado por todos aqueles que desejem falar e escrever bem. Portanto, saber falar

uma língua é conhecer esse sistema e empregá-lo corretamente, pois qualquer desvio é visto como deturpação, índice de ignorância e prejuízo à eficácia e à clareza da comunicação (LUCCHESI, 2015, p. 11).

O mesmo ocorre também com o termo “gramática”, o qual tem sido empregado de forma variada: “No uso convencional, uma gramática é uma descrição ou uma teoria de uma língua, um objeto construído por um linguista” (CHOMSKY, 1994 [1986], p. 39). No entanto, é necessário estabelecer uma associação, correspondente às várias concepções técnicas de línguas, entre as noções de gramática e de gramática universal (GU).

A linguística estrutural e descritiva, bem como a psicologia do comportamento, além de outras abordagens contemporâneas, expressa uma predisposição para reconhecer a língua “como um conjunto de ações, ou produções, ou formas linguísticas (palavras, frases) acompanhadas de significados, ou como um sistema de formas e de eventos linguísticos” (CHOMSKY, 1994 [1986], p. 39). Desse modo, o estruturalismo saussuriano viu a língua, ou *langue*, como um sistema de sons que estava associado a um sistema de conceitos, deixando, de certo modo, de lado a noção de frase, provavelmente a ser incluída no estudo do uso da língua. Já para Bloomfield, de acordo com Chomsky (1994 [1986]), a língua seria o total de enunciados produzidos em uma comunidade linguística:

A variante americana da linguística estrutural-descritiva que foi grandemente influenciada pelas ideias de Bloomfield concentrou-se primariamente no som e na estrutura da palavra, à exceção de várias propostas, nomeadamente as de Zellig Harris, sobre o modo como unidades menores (sintagmas) poderiam ser construídas através de princípios analíticos modelados a partir daqueles que foram introduzidos para a fonologia e para a morfologia. Muitos investigadores adoptam hoje uma posição semelhante à desenvolvida lucidamente por David Lewis, que define uma língua como um emparelhamento de frases e de significados (considerados estes últimos como construções baseadas na teoria dos conjuntos em termos de mundos possíveis) sobre um domínio infinito, em que a língua é “usada por uma população”, quando certas regularidades “na acção ou crença” se verificam entre a população no que respeita à língua, sustentadas por um interesse na comunicação (CHOMSKY, 1994 [1986], p. 39).

Esses conceitos técnicos são categorias de uma “língua externa”, ou língua-E, segundo Chomsky (1994 [1986]), significando que essa é formada separadamente das propriedades da mente/cérebro. Sendo assim, a gramática diz respeito à língua-E:

Sob a mesma rubrica podemos incluir a noção de língua como um conjunto (ou sistema) de acções ou comportamentos de um certo tipo. Deste ponto de vista, uma gramática é uma coleção de afirmações descritivas que dizem respeito à língua-E, a actos linguísticos reais ou potenciais (eventualmente acompanhadas de uma análise do seu contexto de uso ou do seu conteúdo semântico). Em termos técnicos, a gramática pode ser vista como uma função que enumera a língua-E (CHOMSKY, 1994 [1986], p. 39).

A língua-E pode ser considerada como o objeto real de estudo. Desse modo, a gramática seria uma noção derivada, conforme Chomsky (1994 [1986]), cabendo ao linguista a liberdade de escolher a gramática de uma maneira ou de outra, contanto que venha a identificar corretamente a língua-E. Desse modo, conforme Quine (1970), não teria sentido dizer que uma gramática é “correta” em relação à outra, já que ambas são extensionalmente equivalentes e caracterizam a mesma língua-E. Retomando o que foi dito previamente, assim como a língua-E pode ser considerada o objeto real de estudo, para Chomsky há (1978 [1965], p. 107): “Uma gramática pode ser considerada como uma teoria de uma língua”. Por isso mesmo, em se tratando de gramática, não se deve levar em consideração questões acerca de verdade ou falsidade, conforme afirma Chomsky (1994 [1986]).

Em se tratando da GU, esses argumentos são verdadeiros a muitas ou, por que não dizer, a todas as línguas humanas, o que constituiria, de acordo com Chomsky (1994 [1986]), uma associação de condições que satisfazem a todas as línguas-E que funcionam como línguas humanas. Outros argumentos surgiram, no entanto, vindo de encontro a essas afirmações:

Aparentemente alguns negaram a possibilidade de empreendimento da tarefa, como por exemplo Martin Joos, que pôs em destaque aquilo a que chamou o ponto de vista “boasiano”, segundo o qual “as línguas podiam diferir umas das outras sem limites e de maneiras imprevisíveis”, fazendo eco da referência de William Dwight Whitney à “diversidade infinita do comportamento linguístico humano” e à idéia de Edward Sapir de que “a língua é uma atividade humana que varia sem que se lhe possa definir um limite”. (CHOMSKY, 1994 [1986], p. 40).

As afirmações anteriores exprimem um entendimento muito comum na época, mesmo que não tenham sido proclamadas com o intuito de serem interpretadas de forma literal. Caracterizam uma tendência relativista, conforme afirma Chomsky (1994 [1986]), que contribuiu para um descrédito no estudo da GU. Apesar de ser verdadeiro que a língua humana é infinitamente diversa, não há a possibilidade de que ela varie sem que se possa definir um limite:

É uma questão empírica com algum interesse a de saber se a GU permite uma variedade infinita de línguas possíveis (ou uma variedade que é infinita apenas em aspectos estruturalmente triviais, por exemplo, não haver limites quanto ao vocabulário), ou se apenas permite uma diversidade finita (CHOMSKY, 1994 [1986], p. 40).

Desse modo, é possível o desenvolvimento de um conceito técnico de língua, ou seja, língua-E e um conceito associado de gramática e de GU, conforme indica Chomsky (1994 [1986]), que sirva como base para um estudo científico da linguagem.

Partindo da noção de língua-E, a abordagem feita por Otto Jespersen (1935 [1924]), sobre a existência de uma “noção de estrutura” que se faz presente na mente do falante, a qual “é suficientemente definida para o guiar na estruturação das suas próprias frases” (*apud* CHOMSKY, 1994 [1986], p. 41), referindo-se particularmente às “expressões livres” que podem ser novas para o falante e para os outros. Essa noção de estrutura é chamada por Chomsky de língua-I, sendo essa “um elemento que existe na mente da pessoa que conhece a língua, adquirido por quem aprende e usado pelo falante-ouvinte (CHOMSKY, 1994, p. 41).

Ao considerar a língua como língua-I, Chomsky (1994) compreende a gramática como uma teoria da língua-I, sendo objeto a ser investigado. E ainda constata que: “E se, de facto, a tal “noção de estrutura” existe como afirma Jespersen (1924), então surgem questões de verdade e falsidade em relação à gramática, como em relação a qualquer teoria científica” (CHOMSKY, 1994 [1986], p. 41). Dessa forma, retornando, então, a um ponto de vista citado anteriormente:

Conhecer a língua L é uma propriedade de uma pessoa H; uma das tarefas das ciências do cérebro é a de determinar aquilo que existe no cérebro de H que é responsável pela existência dessa propriedade. Sugerimos que dizer que H conhece a língua L significa que a mente/cérebro de H está num certo estado  $E_L$ . Uma das tarefas das ciências do cérebro é a de descobrir os mecanismos que são a realização física do  $E_L$  (CHOMSKY, 1994 [1986], p. 41-42).

Isso conduzirá a uma maneira diferente de abordar a linguagem e a uma concepção que difere, e muito, da natureza da investigação. Desse modo, a GU passa a ser concebida como a teoria das línguas-I humanas, ou seja, “um sistema de condições que deriva do equipamento biológico humano que identifica as línguas-

I que são humanamente acessíveis em condições normais” (CHOMSKY, 1994 [1986], p. 43).

O foco que até então era, de acordo com Chomsky (1994 [1986], p. 43), “o comportamento linguístico real ou potencial e dos produtos desse comportamento” foi substituído pelo “sistema de conhecimento que sustenta o uso e a compreensão da língua”, principalmente no que se refere à capacidade inata que possibilita aos seres humanos alcançar esse conhecimento. Dessa maneira, os estudos da gramática gerativa passam a estabelecer outro objeto de estudo, que migra da língua-E para a língua-I, ou seja, o estudo passa a ser o de uma língua interna, que está reconhecida na mente/cérebro e não mais o de uma língua considerada como um objeto exteriorizado:

Uma gramática gerativa não é um conjunto de asserções acerca de objectos exteriorizados construídos de uma determinada maneira. Pelo contrário, pretende representar exatamente aquilo que uma pessoa sabe quando sabe uma língua, ou seja, aquilo que foi aprendido, porque complementado por princípios inatos. A GU é uma caracterização destes princípios inatos e biologicamente determinados, que constituem uma componente da mente humana – a faculdade da linguagem (CHOMSKY, 1994 [1986], p. 43-44).

Retomamos, então, as questões abordadas anteriormente por Chomsky (1994 [1986], p. 23):

- (1) (i) O que constitui o conhecimento da língua?
- (ii) Como é adquirido o conhecimento da língua?
- (iii) Como é usado o conhecimento da língua?

A partir da mudança de foco referente a essas questões, temos as seguintes respostas dadas:

A resposta a (1i) era a de que o conhecimento da língua consiste no conhecimento de um determinado sistema de regras; a resposta a (1ii) era a de que esse conhecimento deriva de um estado inicial  $E_0$  que converte a experiência num estado estável  $E_E$  que incorpora uma língua-I. A aquisição da linguagem era então uma questão de aumento de provisão de regras de um indivíduo ou de modificação desse sistema, devido ao processamento de novos dados. A questão (1iii) divide-se em duas partes: um “problema de percepção” e um “problema de produção”. O problema de percepção seria resolvido com a construção de um analisador que incorporasse as regras da língua-I, a par com outros elementos: uma certa organização da memória e do acesso (talvez uma estrutura determinística em pilha inversa com uma memória intermediária de uma certa dimensão; vide Marcus, 1980), certas heurísticas, etc. Um analisador não deveria projectar as expressões nas

suas estruturas do modo como estas são associadas pela língua-I. Por exemplo, um analisador deveria falhar no caso das chamadas frases “garden-path” ou no caso de frases que sobrecarregam a memória na passagem da esquerda-para-a-direita (CHOMSKY, 1994 [1986], p. 44).

Em função de o “problema da projeção” ser considerado mais obscuro por Chomsky (1994 [1986]), a discussão acerca dessa questão acaba sendo postergada. Quanto à língua-E, Chomsky (1994 [1986]) considera que essa, que foi objeto de estudo na maioria das gramáticas tradicionais e/ou estruturalistas e até na psicologia comportamental, hoje, para a gramática gerativa, é encarada como um fenômeno secundário, sem muita influência.

Um quadro geral, então, é apresentado por Chomsky (1994 [1986]). Nesse quadro, resume essas questões ao comentar sobre a faculdade da linguagem como um sistema distinto na mente/cérebro, que possui um estado inicial  $E_0$ , e é comum e exclusiva da espécie humana, no que se refere a aspectos essenciais. Além disso, acrescenta:

Dada experiência apropriada, esta faculdade passa de um estado inicial  $E_0$  para um estado final relativamente estável  $E_E$ , que, então, sofre apenas modificações periféricas (como, por exemplo, a aquisição de novo vocabulário). O estado atingido incorpora uma língua-I (é o estado em que se tem ou se conhece uma língua-I particular). A GU é a teoria do  $E_0$ ; as gramáticas particulares são teorias das várias línguas-I. As línguas-I que podem ser atingidas com o  $E_0$  fixo e a experiência linguística variável são as línguas humanas atingíveis, em que “língua” significa agora língua-I (CHOMSKY, 1994 [1986], p. 44-45).

O estado estável possui duas componentes, conforme Chomsky (1994 [1986]), que se distinguem analiticamente, podendo ser fundidas ou entrelaçadas. Há uma componente específica da língua em questão e outra que é o contributo do estado inicial, constituindo na primeira aquilo que é “aprendido”. Chomsky (1994 [1986]) chama a atenção para esse conceito, enfatizando que, quando se trata de dar conta da transição do estado inicial para o estado estável da linguagem, não se sabe ao certo se seria o mais adequado a ser empregado.

Era necessária essa mudança de foco da língua-E para a língua-I, já que, nesse sentido, em pelo menos dois aspectos, o conceito técnico de língua-E é ambíguo, conforme Chomsky (1994 [1986], p. 45): “Em primeiro lugar, tal como foi observado, as línguas (no sentido de língua-E) não são objetos do mundo real, mas sim objetos artificiais, de alguma forma arbitrários, e talvez construções não muito interessantes”. Já o estado estável de conhecimento atingido e o estado inicial, são

considerados por Chomsky (1994 [1986], p. 45), como “elementos reais de mentes/cérebros particulares, aspectos do mundo físico, em que estados mentais e representações são entendidos como sendo de alguma maneira, fisicamente codificados”. A língua-I, como uma constituinte do estado atingido, é diretamente abstraída. Por isso mesmo, as afirmações sobre a língua-I, no que se refere ao estado estável e ao estado inicial  $E_0$ , são questões verdadeiras ou falsas sobre algo que é real e definido, sobre estados reais da mente/cérebro e das suas constituintes:

A GU e as teorias das línguas-I, gramáticas universal e particulares, estão a par de teorias científicas noutros domínios. As teorias das línguas-E, se forem sensíveis, têm um estatuto diferente e mais obscuro, já que não têm um objecto correspondente no mundo real. A linguística, concebida como o estudo da língua-I e do  $E_0$ , torna-se parte da psicologia e, em última análise, da biologia. A linguística está incorporada nas ciências naturais à medida que forem descobertos mecanismos que tenham as propriedades reveladas nestes estudos mais abstractos. De facto, esperar-se-ia que estes estudos fossem um passo necessário para uma investigação séria dos mecanismos. Por outras palavras, a língua-E, embora construída, está mais distante dos mecanismos do que a língua-I, num grau de abstracção mais elevado (CHOMSKY, 1994 [1986], p. 46).

Quando se fala em mudança de foco, remete-se também a uma mudança em relação à noção de língua do senso comum. Quando uma pessoa conhece uma língua, isso não significa que ela conheça “um conjunto infinito de frases ou de pares som-significado tomados extensionalmente, ou um conjunto de acções ou de comportamentos” (CHOMSKY, 1994 [1986], p. 46). A pessoa conhece aquilo que associa de uma determinada forma, no caso o som, ao significado. De certa maneira, o que faz essas partes estarem juntas ou, quem sabe, uma caracterização particular de uma função, é o fato de que “A pessoa tem uma “noção de estrutura” e conhece uma língua-I, tal como é caracterizada pelo linguista” (CHOMSKY, 1994 [1986], p. 46).

Quando se muda a perspectiva do conceito técnico de língua-E para o conceito técnico de língua-I, enquanto objeto de investigação, outros dois aspectos sofrem alterações. Primeiramente, a mudança de perspectiva de estudo de uma construção artificial para o estudo de um objeto real. Por outro lado, “em segundo lugar, no sentido do estudo daquilo que realmente queremos dizer com “língua” ou “conhecimento da língua” num uso informal (fazendo, uma vez mais, abstracção de fatores sociopolíticos e normativos teleológicos)” (CHOMSKY, 1994 [1986], p. 47).

Incidentes cronológicos de publicação e alguns trechos de textos com fins de exposição retirados do contexto, de acordo com Chomsky (1994 [1986]), contribuíram para obscurecer a mudança de conceitos da língua-E para a língua-I, do comportamento e seus produtos para o sistema de conhecimento que intervém no comportamento. O uso do termo “língua”, na literatura gerativista, para denominar língua-E, referindo-se a frases bem formadas, é uma das questões citadas por Chomsky (1994 [1986]), as quais têm contribuído para os mal-entendidos. Desse modo:

O termo “gramática” foi então utilizado com uma ambiguidade sistemática, referindo-se àquilo a que chamamos aqui “língua-I” e também à teoria do linguista acerca da língua-I; o mesmo se pode dizer acerca do termo GU, introduzido mais tarde com a mesma ambiguidade sistemática, referindo-se quer ao  $E_0$  quer à teoria do  $E_0$ . Uma vez que o foco da atenção incidia sobre a língua-I, sendo a língua-E uma construção derivada e altamente artificial, chegamos à situação paradoxal de, nos trabalhos dedicados à língua, o termo “língua” raramente aparecer (CHOMSKY, 1994 [1986], p. 47-48).

Já o termo “língua” deveria ser usado, preferencialmente, para denominar algo mais próximo do sentido intuitivo do uso informal, ou seja: “usar o termo ‘língua’ como um termo mais técnico em vez de ‘gramática (generativa)’ (no sentido de língua-I) e adoptar um termo técnico (talvez “língua-E”) para aquilo a que se chamou ‘língua’”. O termo “gramática (generativa)” teria, então, sido usado para a teoria do linguista acerca da língua(-I)” (CHOMSKY, 1994 [1986], p. 48).

O estudo das línguas formais é enganador, já que a “gramática”, para esse sistema, conforme afirma Chomsky (1994 [1986]), é vista como um conjunto de regras que detalha quais frases são bem formadas. Não se leva em conta, então, “a questão ulterior da escolha correcta da gramática nem a da verdade ou falsidade na escolha entre tais gramáticas” (CHOMSKY, 1994 [1986], p. 49). Assim, volta, então, à conclusão de Quine (1970), citada anteriormente, para reafirmar que não há sentido algum em considerar uma gramática “correcta” em relação à outra, quando são extensionalmente equivalentes:

É de notar que as caracterizações familiares de “língua” como um código ou como um jogo apontam correctamente para a “língua-I” e não para a construção artificial que é a língua-E. Um código não é um conjunto de representações, mas antes um sistema específico de regras que atribui representações codificadas ao par mensagem-representações. Dois códigos podem ser diferentes, embora extensionalmente idênticos nos emparelhamentos mensagem-código que fornecem (CHOMSKY, 1994 [1986], p. 50).

Em se tratando do estudo da linguagem humana, o objeto teórico é questão de recorte e escolha, e não de verdade, para Chomsky (1977), podendo ser investigado de um ponto de vista biológico, assim como de inúmeros outros: o sociolinguístico, o de língua e cultura, o histórico e assim por diante. “Cada uma dessas abordagens define o objeto de sua investigação sob a luz de seus próprios interesses; e, se for racional, cada uma tentará apreender o que puder do que vem das outras abordagens” (CHOMSKY, 1977, p. 3)<sup>1</sup>.

Na próxima seção, o nosso foco volta-se para a Sociolinguística.

## 1.2 A SOCIOLINGUÍSTICA

Nesta seção, após discorrermos sobre a Teoria da Gramática, abordaremos a análise do comportamento social ou do estudo da fala que, de acordo com Labov (2008 [1972]), tem sido negligenciada quando se trata da definição da linguística, o que, de certo modo, é visto por ele como uma questão de conveniência para os formuladores desse estudo. Esses, por causa de uma questão pessoal, preferem trabalhar com um conhecimento próprio, informantes individuais ou com materiais secundários. Essa é uma estratégia bem-sucedida quando se trata da abordagem da estrutura linguística: “Não existe nenhuma razão *a priori* pela qual alguém tenha que entrar na comunidade de fala para buscar dados” (LABOV, 2008 [1972], p. 219). Isso porque é necessário um grande consumo de tempo e esforço que têm que ser justificados. Por outro lado, há, nas últimas cinco décadas, o sucesso da análise linguística abstrata que impossibilitou esse desenvolvimento: “De fato, a limitação de nosso campo de investigação certamente foi útil no desenvolvimento da gramática gerativa – a elaboração de modelos abstratos baseados em nossos julgamentos intuitivos das frases” (LABOV, 2008 [1972], p. 220).

O termo sociolinguística apareceu pela primeira vez em 1952, como título de um trabalho de Currie, segundo Morales (2015 [1993]) e, até os dias atuais, reúne muitos trabalhos publicados e pesquisas divulgadas em livros, artigos, congressos, simpósios, entre outros meios que constituem peças de um *corpus* bibliográfico da

---

<sup>1</sup> Deixaremos para falar sobre a Teoria de Princípios e Parâmetros e o Parâmetro do Sujeito Nulo, um dos assuntos mais discutidos e estudados dentro desta teoria, e que é o objeto de estudo do nosso trabalho, na seção seguinte.

então nova disciplina. Diversos estudos já haviam sido desenvolvidos por Meillet (1866-1936), Baktin (1895-1975) e outros membros do Círculo Linguístico de Praga, antes dos anos 1960, conforme Bortoni-Ricardo (2014, p. 11) com teorias de caráter visivelmente sociolinguístico. Desde o seu nascimento, falar sobre sociolinguística exige, em primeiro lugar, falar o que se entende por ela:

Lo que entonces se definía como “relaciones entre lengua y sociedad” y más recientemente como “lengua y contexto sociocultural” (Newmeyer, ed., 1988) ha dado pie a que investigaciones de variado tipo, alcance y objetivos, hayan sido clasificadas como sociolingüísticas, aun – sorprendentemente – aquellas que nada o poco tienen que ver con la lingüística (MORALES, 2015 [1993], p. 7).<sup>2</sup>

Mesmo que também tenham sido produzidos estudos em outros lugares, os pilares mais importantes, as suas bases, foram levantadas, conforme afirma Morales (2015 [1993]), nos Estados Unidos, na parte francesa do Canadá e na Inglaterra, ainda que não houvesse, à época, uma diferenciação precisa entre a sociolinguística propriamente dita e outras disciplinas, como sociologia da linguagem, etnografia da comunicação, estudos do bilinguismo, organização do discurso e análise da conversação, a maior parte das pesquisas tem procedência em tais centros (Cf. Morales, 2015 [1993], p. 8). Já em relação à sociolinguística desenvolvida nos Estados Unidos e no Canadá na década de 1960, segundo Silva-Corvalán (1988), preocupa-se com o contexto externo aos fatos linguísticos e teve como objeto de estudo a variação linguística, o modo como ela se inter-relaciona no meio social e o papel desempenhado por ela nos processos de mudança linguística.

Sobre a distinção entre a sociolinguística e a sociologia da linguagem, Silva-Corvalán (1988), ao se referir à metodologia empregada nas pesquisas sociolinguísticas, afirma que a sociolinguística estuda fenômenos linguísticos no contexto social, e a sociologia da linguagem estuda os fenômenos sociais relacionados com o uso de uma língua. Quanto à dialetologia, cujos trabalhos iniciais sobre a linguagem são mais antigos que os estudos sociolinguísticos, Silva-Corvalán (1988) comenta o fato dessa, posteriormente, ter introduzido em suas pesquisas

---

<sup>2</sup> O que então se definia como “relações entre língua e sociedade” e mais recentemente como “língua e contexto sociocultural” (Newmeyer, ed., 1988) deu pé a que investigações de variado tipo, alcance e objetivos, fossem classificadas como sociolinguísticas, ainda – surpreendentemente – aquelas que nada ou pouco têm que ver com a linguística (tradução nossa).

alguns elementos da sociolinguística moderna, no que se refere à observação dos fatos sociais em relação às variáveis.

A sociolinguística possui uma relação direta com fatores sociais, conforme Silva-Corvalán (1988), o que inclui uma análise política, econômica, social e geográfica. Isso abarca também fatores sociais que têm uma relação com o indivíduo e a sua convivência em sociedade como idade, sexo, nível de educação, etnia, a história e a própria situação de interação no momento imediato. A Sociolinguística também é definida como uma subárea da Linguística, segundo Mollica (2004) que, por seu turno estuda a língua em uso em uma determinada comunidade de fala, utilizando-se de uma investigação específica que compara aspectos linguísticos e sociais. A sociolinguística é uma ciência interdisciplinar e de caráter heterogêneo, conforme Mollica (2004) e Bortoni-Ricardo (2014), tendo como foco os empregos linguísticos concretos. Já Coelho *et al.* (2015, p.13) afirmam que o termo sociolinguística nos remete tanto à palavra *linguística*, quanto à palavra *social*, já que de fato, “A Sociolinguística é uma área da Linguística que estuda a relação entre a língua que falamos e a sociedade em que vivemos”.

Através de uma iniciativa de William Bright, de acordo com Calvet (2002), no ano de 1964, foram reunidos para uma conferência sobre a sociolinguística, em Los Angeles, 25 pesquisadores. A maioria composta por americanos, quinze no total, e mais dois participantes da Iugoslávia, mas que se encontravam temporariamente na UCLA, a universidade que organizou o evento, sendo que 13 desses participantes apresentaram comunicações sobre diversos temas:

Os temas abordados eram variados: a etnologia da variação linguística (Gumperz), a planificação linguística (Haugen), a hipercorreção como fator de variação (Labov), as línguas veiculares (Samarin, Kelley), o desenvolvimento de sistemas de escrita (Sjoberg), a equação de situações sociolinguísticas dos Estados (Ferguson)... e os referenciais teóricos não eram menos variados (CALVET, 2002 [1993], p. 29).

Bright (1974 [1966]) fez uma tentativa de sintetizar as diferentes contribuições dessa conferência, indicando a dificuldade de se definir com precisão a sociolinguística, cujos estudos se referem às relações entre linguagem e sociedade. Diante dessa afirmação, considerada por ele extremamente vaga, enfatiza que a sociolinguística vai diferir de outras inquietações anteriores que dizem respeito às relações entre língua e sociedade “pelo fato de que, seguindo novas perspectivas da

própria linguística, considera tanto a língua quanto a sociedade como sendo uma estrutura e não uma coleção de itens” (BRIGHT, 1974 [1966], p. 17). Seria tarefa, então, da sociolinguística, a comprovação da covariação sistemática presentes nas variações linguística e social, constatando a relação causal existente em uma ou outra direção.

Os sociolinguistas se afastaram decisivamente da tendência linguística que tratava a estrutura das línguas como totalmente uniformes, homogêneas ou monolíticas. Nesse sentido, Bright (1974 [1966], p. 18) afirma que: “As diferenças encontradas nos hábitos de fala de uma comunidade eram encobertas como ‘variação livre”. Sendo assim, esclarece ainda que a sociolinguística tem como principal tarefa indicar que a variação ou a diversidade não é livre, mas tem uma relação direta com as diferenças sociais de forma sistemática. A partir desse ponto de vista, Bright (1974 [1966], p. 18) assevera que “É precisamente a DIVERSIDADE linguística o objeto de estudo da sociolinguística”. O autor complementa afirmando que também considera essa afirmação incompleta para dar conta do enorme conjunto de estudos sociolinguísticos possíveis. Por isso mesmo, diz ser necessário tentar identificar as “dimensões” da sociolinguística, que envolvem diversas correntes independentes, para uma maior aproximação da descrição desses estudos. Haveria, então, de acordo com Bright (1974 [1966]), para cada interseção de duas ou mais dessas dimensões, um objeto de estudo da sociolinguística apresenta-se, então, sete dessas dimensões.

Ao considerar a diversidade como conceito-chave, Bright (1974 [1966]) relaciona o condicionamento da diversidade linguística a um conjunto de dimensões considerado mais importante. Este, por sua vez, se refere a diversos fatores definidos socialmente e que se acredita que estejam relacionados à diversidade linguística: a identidade social do falante, a identidade social do ouvinte e o contexto. O primeiro desses três fatores citados, e que diz respeito à identidade social do “emissor” ou falante, segundo Bright (1974 [1966]), relaciona-se à diferença da fala conforme a estratificação social, observada claramente nos “dialetos de classe”. Já o segundo fator, refere-se à identidade social do “receptor” ou ouvinte, quando são usados vocabulários específicos que indicam respeito a superiores, por exemplo. Uma terceira dimensão condicionadora, também citada por Bright (1974 [1966]), refere-se ao contexto e tem relação com a identidade dos indivíduos de um mesmo grupo social e tudo o que é relevante na comunidade linguística da qual esses

indivíduos fazem parte. Ressalta ainda não haver exclusividade em relação a cada uma dessas três dimensões e que elas podem, obviamente, se interseccionar condicionando um tipo específico de comportamento linguístico.

As outras dimensões da sociolinguística, de acordo com Bright (1974 [1966]), não são muito baseadas no comportamento linguístico evidente, mas tem a ver com a finalidade e os objetivos do pesquisador. Em relação às quatro restantes, há ainda a oposição sincronia/diacronia, que está presente também na pesquisa sociolinguística, assim como em outras áreas. Outra dimensão diz respeito ao modo como as pessoas usam a língua e a crença acerca de seu próprio comportamento linguístico e do comportamento dos outros.

A extensão da diversidade, apresentada por Bright (1974 [1966]), incluiria três classificações. A primeira delas foi rotulada como pluridialetoal e implica o uso de variedades de uma mesma língua socialmente condicionadas em uma única sociedade ou nação. Outra classificação denominada plurilíngue refere-se ao uso de diferentes línguas dentro de uma única sociedade ou nação. A terceira classificação chamada de plurisocietal estuda línguas distintas que são faladas em diferentes sociedades. A última dimensão apresentada é a da aplicação, inerente às descrições da diversidade sociolinguística, dividida em três categorias. A primeira delas abrange o diagnóstico da estrutura social, através da utilização de dados sociolinguísticos. A segunda aplicação representa o interesse do linguista no estudo do fator sócio-histórico. Já o terceiro tipo de aplicação, diz respeito ao planejamento linguístico em uma sociedade organizada em que são encontradas uma diversidade de dialetos ou línguas.

A Sociolinguística, então, enquanto ciência autônoma e interdisciplinar tem as suas bases iniciais em meados do século XX, conforme Bortoni-Ricardo (2014), e considera o contexto sociocultural e a comunidade de fala em seus estudos, sem separar a fala dos indivíduos das condições em que é produzida, considerando essa associação extremamente importante. Os pioneiros da Sociolinguística nos Estados Unidos, liderados por William Labov, nos momentos iniciais em que se consolidavam as bases dessa ciência, desenvolveram pesquisas voltadas, de acordo com Bortoni-Ricardo (2014), que descreviam a variação dos fenômenos em processo de mudança inerentes à língua, o que foi ampliado posteriormente também para outras dimensões da linguagem humana.

A seguir, enfocamos a visão de Labov sobre o estudo da língua dentro do contexto social da comunidade de fala, os problemas na relação entre dados e teoria, dentre outros aspectos da teoria.

### 1.2.1 Labov e a Sociolinguística

O procedimento que aborda a pesquisa linguística e se concentra na língua em uso dentro da comunidade de fala, visando uma teoria linguística que seja adequada para poder analisar esses dados, para Labov (2008 [1972]), tem sido denominado de “sociolinguística”, sendo que acrescenta, ainda, que esse termo tem sido usado, redundantemente, de forma estranha e enganosa:

A língua é uma forma de comportamento social: declarações neste sentido podem ser encontradas em qualquer texto introdutório. Crianças mantidas em isolamento não usam a língua; ela é usada por seres humanos num contexto social, comunicando suas necessidades, ideias e emoções uns aos outros. Os monólogos egocêntricos das crianças revelam ser desenvolvimentos secundários derivados do uso social da língua (Vigotsky 1962:19) e pouquíssimas pessoas passam muito tempo falando consigo mesmas. É questionável se frases que nada comunicam a ninguém façam parte da língua. De que maneira, então, a “sociolinguística” pode ser considerada algo separado da “linguística”? (LABOV, 2008 [1972], p. 215).

Por isso mesmo, ao abordar a questão do estudo da estrutura e da evolução da língua dentro do contexto social da comunidade de fala, Labov (2008 [1972]) argumenta que:

Se não houvesse necessidade de contrastar este trabalho com o estudo da língua fora de seu contexto social, eu preferiria dizer que se trata simplesmente de *linguística*. É relevante, portanto, indagar por que deveria haver a necessidade de uma nova abordagem da linguística com uma base social mais ampla. Parece bastante natural que o dado básico para qualquer forma de linguística geral seja a língua tal como usada por falantes nativos comunicando-se uns com os outros na vida diária (LABOV, 2008 [1972], p. 216-217).

Mesmo com as dificuldades e problemas que ocorrem ao lidar com a fala, não se pode descartar o estudo da língua de seu contexto social. Por isso mesmo, em vários momentos, os linguistas buscaram comprovação a partir do estudo da fala. Dentro desse contexto, Labov (2008 [1972]) também apresenta quatro dificuldades que são citadas como tendo produzido fortes efeitos sobre a prática da linguística: (i)

a agramaticalidade da fala; (ii) variação na fala e na comunidade de fala; (iii) dificuldades de ouvir e gravar; (iv) a raridade das formas sintáticas.

Sobre a *agramaticalidade da fala*, apesar de linguistas da escola bloomfieldiana terem declarado, em um determinado momento, que os falantes nativos não cometiam erros, para Labov (2008 [1972]) o contrário prevalece nos dias de hoje, já que se sabe que a fala é repleta de formas agramaticais. Isso ocorre porque “as dificuldades de desempenho se colocam no caminho da plena manifestação da competência do falante” (LABOV, 2008 [1972], p. 220). Além disso, é importante salientar que “Acredita-se em geral que um *corpus* extraído da língua falada não constitui boa evidência, já que conterá vários exemplos de frases malformadas que os próprios falantes condenam e mudam quando sua atenção é chamada para elas” (LABOV, 2008 [1972], p. 220).

Quanto à questão da *variação na fala e na comunidade de fala*, Labov (2008 [1972]) diz ser natural uma língua possuir diversas opções para se dizer a “mesma” coisa. Há palavras que parecem ter referentes semelhantes, enquanto outras podem ter mais de uma pronúncia ou mais de uma opção sintática. Nesses casos, o problema é a decisão referente ao lugar da variação na estrutura linguística. São oferecidas claramente, então, duas opções através da análise contemporânea, conforme Labov (2008 [1972], p. 221):

- (1) diz-se que as variantes pertencem a dois sistemas diferentes, e que a alternância é um exemplo de “mistura dialetal” ou “alternância de código” [*code-switching*];
- (2) diz-se que as variantes se encontram em “variação livre” dentro do mesmo sistema, e a seleção se encontra abaixo do nível da estrutura linguística.

As duas abordagens deixam a variação fora do sistema que está sendo estudado. Labov (2008 [1972], p. 221) complementa, afirmando que:

Existem, é claro, diversos casos aos quais cabe apropriadamente um ou outro desses rótulos. Mas para demonstrar que temos um verdadeiro caso de alternância de código, é necessário mostrar que o falante se move de um conjunto consistente de regras coocorrentes para outro; para demonstrar “a variação livre”, é preciso mostrar que ele não se moveu de jeito nenhum. É raro que qualquer dessas alegações seja estabelecida empiricamente (LABOV, 2008 [1972], p. 221).

Ao falar sobre as *dificuldades de ouvir e gravar*, Labov (2008 [1972]) comenta que quase sempre as gravações de fala observada em uso real possuem uma

qualidade muito deficiente. Quando efetuado um trabalho de campo, ruídos no local onde estão sendo feitas as gravações na sala, ou mesmo barulhos externos na rua, por exemplo, prejudicam interferindo na qualidade fonética dos dados<sup>3</sup>. Então, poderia se optar em gravar em uma sala à prova de som, em condições excelentes, as melhores possíveis. No entanto, ao ser levado para um local ideal, para realizar essas gravações, o informante tenderá a aproximar a sua fala de propriedades formais, o que justamente se quer evitar:

O problema fundamental é que a maioria dos sinais linguísticos são apoiados por uma grande quantidade de sinais redundantes, e é raro que qualquer um deles transmita uma carga pesada de significado; não é essencial, para a mensagem global, que os ouvintes percebam todos os sinais. No entanto, para gravar este item na forma plena, o linguista gostaria de ouvi-lo do modo mais claro, como se fosse o único meio de assinalar a mensagem. Daí decorreria, portanto, que as formas direcionadas obtidas no laboratório oferecem a indicação mais clara do sistema subjacente (LABOV, 2008 [1972], p. 223).

O quarto e último item sobre os problemas em lidar com a fala, de acordo com Labov (2008 [1972]), diz respeito à *raridade das formas sintáticas* e refere-se às dificuldades na obtenção de dados específicos, ou seja, de formas específicas que são o enfoque do linguista em uma determinada investigação. Nesse caso, o investigador terá que lidar com a limitação de dados. De acordo com Labov (2008 [1972]), essa inadequação do *corpus* médio se torna mais evidente no estudo da sintaxe. Toda tentativa de especificação de regras sintáticas implicará decisivamente em formas que não poderão ser ouvidas numa investigação limitada. Isso leva à concentração na *langue* ou “competência”, como motivação básica, já que essas dificuldades contribuem para a exclusão de outros dados.

Uma vez que se tem feito considerável progresso no estudo abstrato da *langue*, e dadas tais dificuldades de trabalho num contexto natural, não surpreende que a linguística tenha se afastado decididamente da comunidade de fala. Mas também existem desvantagens no estudo abstrato da língua. Algumas de suas limitações têm se tornado penosamente conhecidas; as dificuldades de desenvolver a teoria linguística com essa base de dados limitada talvez sejam maiores do que as esboçadas acima para o estudo da comunidade de fala (LABOV, 2008 [1972], p. 224).

---

<sup>3</sup> Os avanços tecnológicos têm permitido, atualmente, que problemas deste tipo sejam dirimidos, impedindo que barulhos externos interfiram na captação da voz do informante, através da utilização de aparelhos de gravação de voz digital e com dispositivo com essa finalidade. Importante salientar que essa questão, na realidade, não é um problema linguístico. É análogo ao problema de os textos não terem sobrevivido para a linguística histórica.

Há um compartilhamento de um conjunto de padrões normativos que são comuns, segundo Labov (2008 [1972]), entre os membros de uma comunidade de fala. Isso ocorre mesmo quando é encontrada uma variação bastante estratificada na fala real. No entanto, “essa uniformidade, nos julgamentos intuitivos é característica somente de variáveis sociolinguísticas bem desenvolvidas que têm recebido correção social explícita” (LABOV, 2008 [1972], p. 225-226). Sendo assim, a maior parte das regras linguísticas se encontra bem abaixo do nível de correção, não tendo, portanto, normas sociais explícitas associadas a elas. Quanto mais os linguistas se envolvem e se aprofundam em questões teóricas, maior a chance de um distanciamento de suas intuições acerca das pessoas comuns e da realidade da língua utilizada por essas em sua vida cotidiana, de acordo com Labov (2008 [1972], p. 233): “Os linguistas não podem continuar a produzir ao mesmo tempo dados e teoria”. Desse modo, é necessário um maior desenvolvimento das técnicas de investigação, mas é improvável que se obtenha julgamentos consistentes e homogêneos dos falantes nativos sobre essas questões:

A variação nos julgamentos sintáticos pode ser estudada com proveito, e as séries implicacionais dentro deles podem ser analisadas para se decidir a forma das regras (Elliott, Legum & Thomson 1969). Mas já ficou evidente que a busca pela homogeneidade em julgamentos intuitivos leva ao fracasso. Quando se aceitar isso, desaparecerá a motivação mais forte para confinar a análise linguística a tais julgamentos. De diversas maneiras, a intuição é menos regular e mais difícil de interpretar do que a fala. Se quisermos fazer bom uso das declarações dos falantes sobre a língua, temos que interpretá-las à luz de produções inconscientes, sem reflexão. Sem esse controle, lidamos com dados realmente muito duvidosos – sem nenhuma relação clara com o processo comunicativo que reconhecemos ser a própria língua (LABOV, 2008 [1972], p. 233).

Sobre essa afirmação, entendemos que, ao buscar certo distanciamento da teoria e focar a sua pesquisa nos dados encontrados, o investigador tem mais liberdade para analisar os resultados. Sabemos também da importância da teoria na investigação e, por isso mesmo, é possível um diálogo entre esses dois métodos, buscando conciliar o levantamento dos dados com o estudo de teorias que servirão de base para a pesquisa. Por outro lado, o linguista não pode deixar de lado o levantamento de dados e dar prioridade às teorias, pois aqueles servirão para a comprovação ou não destas.

Para seguir adiante no estudo da língua é necessário também, conforme afirma Labov (2008 [1972]), ser capaz de trabalhar nesse nível de abstração. É difícil

não concluir que o objeto da linguística deve ser o instrumento de comunicação que é usado pela comunidade de fala: “Se não estivermos falando *desta* língua, há qualquer coisa de trivial em nosso procedimento. Por diversas razões, este tipo de língua tem sido o objeto de interesse mais difícil para a linguística” (LABOV, 2008 [1972], p. 220).

Ainda sobre os problemas enfrentados pelos linguistas ao lidar com os dados diretamente, essa questão não é exclusividade da área, mas de todas as ciências sociais, de um modo geral, conforme afirma Labov (2008 [1972], p. 235): “Garfinckel (1967) demonstrou que existe em todo campo de pesquisa uma lacuna inevitável entre os dados brutos, tal como ocorrem, e os protocolos que registram tais dados como matéria-prima para a elaboração teórica”. Há, na sociolinguística, vários tipos de dados que permitem a coleta de informações sobre a língua no uso real, como censos, questionários, excertos de peças teatrais e romances, testes psicológicos, relatórios etnográficos de normas comunitárias, dentre outros. Mesmo sendo esses estudos perspicazes e produtivos, isso não indica que seja possível chegar mais próximo dos dados fundamentais da língua em uso. Há muitas questões sobre esse problema que não serão respondidas:

Qual é a relação entre o estereótipo criado pelo romancista e o comportamento linguístico das pessoas em questão? Qual é a conexão entre testes de associação de palavras e a semântica da língua natural? Como descobrir quando um falante usa *tu*, se tudo o que temos é seu depoimento pessoal? Ou como saber quando ele fala francês somente perguntando-lhe isto diretamente? Qual a relação entre as normas que o antropólogo relata e a prática dos membros na conformação a essas normas? (LABOV, 2008 [1972], p. 235).

Entre os dados e o relatório do linguista, como indica Labov (2008 [1972]), há vários procedimentos de percepção, recordação, seleção, interpretação e tradução, que estariam implícitos em seu trabalho. Ele destaca a afirmação de Garfinckel, o qual mencionou o procedimento de codificação e transcrição. Esses, ao transformar os dados, exibem um irreduzível resquício de ações derivadas do senso comum, que não é possível sintetizar em regras:

Para lidar com a *língua*, temos de olhar para os dados da fala cotidiana o mais perto e diretamente possível, e caracterizar seu relacionamento com as teorias gramaticais do modo mais acurado que pudermos, corrigindo e adequando a teoria para que ela se ajuste ao objeto visado. Podemos, assim, reexaminar os métodos que temos empregado, numa investigação

que fará crescer enormemente nosso entendimento do objeto que estamos estudando (LABOV, 2008 [1972], p. 235-236).

No entanto, apesar das críticas feitas aos métodos linguísticos convencionais, Labov (2008 [1972], p. 236) salienta não se tratar de uma sugestão para que eles sejam deixados de lado: “A elicitación formal de paradigmas, o exame dos julgamentos intuitivos, o estudo de textos literários, a experimentação em laboratório, os questionários sobre uso linguístico – são todos modos de investigação importantes e valiosos”. Ele acrescenta ainda que qualquer um que tenha como objetivo empreender uma análise linguística significativa tem de levar em conta os dois primeiros procedimentos. Sendo assim, a realização da transcrição fonética de uma língua ou de palavras que sejam desconhecidas vai muito além da capacidade de quem está realizando esse procedimento. O ouvido humano não é suficiente para verificar a qualidade de forma absoluta de sons isolados. No entanto, a obtenção de um conhecimento da sintaxe e dos morfemas, faz com que o ouvido se torne um instrumento excelente, possibilitando julgar as diversas possibilidades dentre as que estão sendo realizadas:

Em sintaxe, nossas primeiras análises de dada forma são relativamente superficiais; mas, quando se nota diversas relações com outras estruturas frasais, começa a aparecer um campo rico de estruturas subjacentes possíveis. Há aqui um segundo *paradoxo cumulativo*: *quanto mais se conhece uma língua, mais se pode descobrir sobre ela* (LABOV, 2008 [1972], p. 236).

Sendo assim, é importante salientar a contribuição da sociolinguística para os estudos linguísticos modernos, principalmente no que diz respeito ao estudo da estrutura e da evolução da língua dentro do contexto social da comunidade de fala. Apesar de todas as dificuldades e problemas que são peculiares ao lidar com a fala, não se pode descartar o estudo da língua de seu contexto social. Por outro lado, além da coleta de dados e a sua análise, é necessário entender a relevância da teoria na investigação, já que, como foi afirmado acima por Labov (2008 [1972]), é possível que esses dois métodos dialoguem entre si, para que se possa agregar o levantamento dos dados com o estudo de teorias que servirão de base para a pesquisa. Sobre essa questão falaremos a seguir, ao abordarmos a sociolinguística paramétrica.

### 1.2.2 Sociolinguística Paramétrica

Chomsky (1966) impôs limitações aos dados de *input*, que levaram à convicção de que a *teoria não é determinada pelos dados*. Isso implica a afirmação de que haverá sempre diversas análises possíveis para cada conjunto de dados, o que necessitaria de medidas de avaliação internas para que se possa escolher entre elas. Por outro lado, Labov (2008 [1972]) assume posição contrária, argumentando que o montante de dados disponíveis se expande e muito, oferecendo formas e meios para uma decisão acerca da possibilidade das várias análises estarem corretas, quando se faz uso de um estudo direto da língua em seu contexto social: “Em nossas operações preliminares sobre os dados iniciais, considerações de simplicidade sempre terão lugar; mas encontrada a correta linha de ataque, é possível provar se a hipótese simples inicial é a correta” (LABOV, 2008 [1972], p. 237).

A apreciação dos dois modelos teóricos sugeriu uma incompatibilidade entre a gramática gerativa e a sociolinguística, em um momento anterior, o que tem provocado um debate, através de estudos mais recentes que propõem um diálogo entre esses dois modelos, através da Sociolinguística Paramétrica, conforme veremos adiante. Neste trabalho, estamos propondo esse diálogo, ao utilizarmos elementos da sociolinguística variacionista e da gramática gerativa na análise dos dados.

Ao usar o termo sociolinguística paramétrica em um artigo intitulado *Por uma sociolinguística românica “paramétrica”: fonologia e sintaxe*, Tarallo (1987) objetivava propor que se fizesse uma leitura comparativa entre as pesquisas variacionistas elaboradas acerca das línguas românicas, conforme o critério conquistado através do modelo teórico de Princípios e Parâmetros (P&P), estabelecido por Chomsky (1988 [1981]), no domínio da sintaxe das línguas. Seria então a proposta de uma leitura, com base na Teoria de P&P, dos resultados obtidos através de seus estudos dentro da Teoria da Variação e Mudança. Desse modo, de acordo com Duarte (2016), não era mais aceitável a negação da compatibilidade do empirismo da sociolinguística com a teoria formal de P&P. Também denominada de socioparamétrica, a sociolinguística paramétrica consiste, de acordo com Duarte (2016, p. 34), na “aplicação do modelo da Sociolinguística (ou Teoria da Variação e

Mudança Linguística, como é também conhecido), utilizando como componente gramatical alguns pressupostos da Teoria de P&P”.

Também para Lucchesi (2011) é possível sim que haja um diálogo entre esses dois modelos. De acordo com incomensurabilidade, formulada por Thomas Kuhn (1970), “a irredutibilidade entre os dois modelos decorreria da ausência de um corpo de conceitos comuns, através dos quais se poderia problematizar a construção do objeto de estudo” (LUCCHESI, 2011, p. 228). No entanto, a reunião da teoria da variação e da teoria da gramática, denominada como variação paramétrica, proposta por Tarallo (1987), sofreu críticas não favoráveis, como a de Borges Neto (2004), que defendia a irredutibilidade entre os dois modelos teóricos. Para Lucchesi (2011), é possível estabelecer um diálogo entre esses dois modelos teóricos, já que é através deles que se manifestam as soluções encontradas pelos sociolinguistas e gerativistas para suplantar algumas contradições idealizadas por Saussure e que atravessaram a era estruturalista da linguística moderna:

Um dos principais problemas postos por Saussure está em situar empiricamente o objeto de estudo da Linguística. A contradição radica na localização da língua, sistema unitário e invariável, na coletividade, o plano *par excellence* da variação e da heterogeneidade. Enquanto a Sociolinguística postulará que o objetivo de estudo da Linguística, a *gramática da comunidade de fala*, só poderá ser pensada nos termos de um sistema heterogêneo e variável, os gerativistas se refugiarão no *idioleto*, na gramática do falante-ouvinte individual, para preservar a homogeneidade e a invariância do seu objeto de estudo (LUCCHESI, 2011, p. 228-229).

Nesse caso, o problema está em tentar conciliar a variação linguística, principalmente a mudança e o funcionamento da língua. De acordo com Kroch (2001), o modo como as mudanças se espalham e por que isso acontece ainda não é muito bem entendido, ainda que tenham sido propostos modelos referentes a esse processo. Um desses modelos, apresentado por Niyogi e Berwick (1997), refere-se a sistemas dinâmicos em que os aprendizes, neste caso as crianças, nem sempre convergem para a língua-alvo a que estão expostas. Sendo assim, quando: *“the evidence for a given parameter setting becomes weak enough, some learners will, due to random effects, not be exposed to enough data to set the parameter correctly”*

(KROCH, 2001, p. 721)<sup>4</sup>. Essa situação resultará em uma população mista de falantes, sendo que alguns terão a fixação paramétrica antiga e outros, a nova. A próxima geração de aprendizes, então, nessa população mista, terá em média, uma exposição menor aos dados necessários para fixar o parâmetro do modo antigo. Ainda de acordo com Niyogi e Berwick (1997), com a expansão dessa população, em muitos casos, vai acontecer mudança da gramática original para a nova ao longo de uma trajetória em uma curva em “S”.

O modelo de Niyogi e Berwick (1997), além de seu caráter hipotético, encontra uma dificuldade ao presumir que os conjuntos de parâmetros em competição se localizam em falantes diferentes, de modo que o elemento quantitativo na mudança sintática não estaria localizado no indivíduo e sim na população. Desse modo, os estudos empíricos que revelam a natureza gradual da mudança não trazem dados consistentes com o modelo apresentado por Niyogi e Berwick (1997) a esse respeito.

Para definirmos essa variação, segundo Kroch (2001), é fundamental que a diglossia sintática seja permitida individualmente entre os autores, durante o período de mudança como uma situação normal:

Again, this conclusion is a logical consequence of the general assumptions of generative theory regarding the categorical nature of grammatical parameters. Furthermore, it is necessary to allow a description of individual speakers under which they have a propensity to choose between their diglossic grammars at a characteristic average rate. This rate, moreover, seems to characterize entire speech communities, and it is what changes over time as one of the grammars slowly drives the other one out of use (KROCH, 2001, p. 722)<sup>5</sup>.

Enquanto esse modo de pensar é, evidentemente, adotado pela sociolinguística, de acordo com Kroch (2001), há uma objeção dos gerativistas em relação a ele. Ainda para Kroch (2001), não há como ter dúvidas sobre o fato dos seres humanos, assim como todos os animais, seguirem as frequências dos eventos

---

<sup>4</sup> a evidência para a fixação de um dado parâmetro se torna fraca o suficiente, alguns aprendizes, devido a fatores aleatórios, não vão ser expostos a dados suficientes para fixar o parâmetro corretamente (tradução de Silvia Regina Cavalcante - UNICAMP).

<sup>5</sup> De novo, essa conclusão é uma consequência lógica dos pressupostos gerais da teoria gerativa que concernem a natureza categorial dos parâmetros gramaticais. Além disso, é necessário permitir uma descrição dos falantes individuais sob a qual eles têm a propensão para escolher entre suas gramáticas diglósicas numa razão média característica. Essa razão, além do mais, parece caracterizar comunidades de fala inteiras, e é isso que muda ao longo do tempo enquanto uma das gramáticas lentamente substitui a outra (tradução de Silvia Regina Cavalcante - UNICAMP).

no ambiente em que convivem. Isso incluiria, no caso dos seres humanos, os eventos linguísticos. O conflito gerado entre a sociolinguística e os gerativistas, nesse caso, de acordo com Kroch (2001), diz respeito ao fato dos sociolinguistas adotarem a proposta de que as probabilidades devam ser integradas às gramáticas, o que não é consistente com o paradigma gerativo:

It is not necessary, however, to make this last move in order to relate variation in usage by individuals to syntactic change. Once a community becomes diglossic with respect to a given parameter setting, every speaker will learn both settings. The choice of which criterion of well-formedness to apply in the production of a given utterance is one that falls in the domain of performance and so is not an issue for grammatical theory. How learners acquire diglossic competence is, of course, an important issue for language acquisition, but there is no doubt that they do (KROCH, 2001, p. 722).<sup>6</sup>

Sendo assim, é de se esperar que os membros de uma comunidade linguística, conforme afirma Kroch (2001), convirjam mais ou menos na mesma frequência de uso, utilizando-se de um conjunto de variantes disponíveis, sendo que essa frequência, tampouco, deva variar ao longo do tempo.

Um dos pilares da sociolinguística e que perpassa toda a sua teorização é a afirmação de que a mudança ocorre durante o funcionamento da língua, ou seja, a língua funciona enquanto muda. Reconhecer esse fato é negar que a mudança e a variação têm um caráter disfuncional, o que foi sugerido antes pela visão estruturalista de língua.

No entanto, a Sociolinguística mostrou-se incapaz de compreender o sistema linguístico como um todo, ao focar a variação e a mudança linguística, e, ao mesmo tempo, não conceber uma teoria para o funcionamento da língua. A Gramática Gerativa, por outro lado, segundo Lucchesi (2011), apesar dos progressos que tem experimentado na formulação do sistema da aquisição da língua humana, se depara com a fragilidade de procedimentos empíricos, que se baseiam na sua

---

<sup>6</sup> Não é necessário, entretanto, dar esse último passo a fim de relacionar a variação no uso pelos indivíduos com a mudança sintática. Uma vez que uma comunidade se torna diglósica com relação a uma dada fixação paramétrica, todo falante vai aprender ambos os parâmetros. A escolha de qual critério de boa-formação deva ser aplicado na produção de uma dada instância de fala recai no domínio da performance e, portanto, não é uma questão para a teoria gramatical. Obviamente, uma questão importante para a aquisição da linguagem é como os aprendizes adquirem uma competência diglósica, mas não há dúvida que eles a adquiram (tradução de Silvia Regina Cavalcante - UNICAMP).

fundamentação em juízos subjetivos e instáveis da gramaticalidade. No entanto, essa afirmação é questionada, pois há também estudos gerativistas com base em *corpus*, principalmente os estudos diacrônicos.

A pesquisa de materiais históricos pode trazer importantes contribuições, apresentando dados que, de acordo com Kroch (1989), certamente estariam ausentes em estudos sincrônicos da linguagem. Desse modo, contribui para a história da linguística geral, oferecendo informações sobre a mudança da língua em um determinado curso de tempo. Conforme Kroch (1989), ao conhecermos esse processo, haveria a possibilidade de se compreender como as línguas mudam ao longo do tempo e, compreendendo esse processo de mudança, seria possível aprendermos mais sobre o modo como se organizam:

In addition, since the features of any language at a given point in time are the result of a complex interweaving of general principles of language and particular historical developments, knowledge of the historical process by which a language has reached a given state may be important to the proper assignment of responsibility to historical and general factors and so to the proper formulation of linguistic theory (Kroch, 1989, p. 200)<sup>7</sup>.

No que se refere às definições da mudança linguística e a aparente incompatibilidade da diferente postura teórica entre Lightfoot (1989) e Labov (2013 [1975]), é possível enxergar uma aproximação entre os dois modelos teóricos, no sentido de uma utilização de ambos em um trabalho investigativo. Enquanto para Lightfoot (1989) a busca das propriedades invariantes da gramática seria o mais importante no estudo da mudança, viabilizando uma descrição das gramáticas possíveis, para Labov (2013 [1975]), o que importa não é apenas a questão da origem, mas principalmente a implementação das mudanças linguísticas. Segundo Tarallo (1991), apesar de não haver dúvidas sobre essas diferentes posturas, de certa forma, radicalmente opostas, entretanto esta mesma incompatibilidade teórica pode ser minimizada ao se diferenciarem os níveis de origem e de propagação da mudança linguística:

---

<sup>7</sup> Além disso, uma vez que as características de qualquer língua em um determinado momento no tempo são o resultado de um complexo entrelaçamento dos princípios gerais da linguagem e dos desenvolvimentos históricos particulares, o conhecimento do processo histórico pelo qual uma língua chegou a um determinado estado pode ser importante para a adequada atribuição da responsabilidade a fatores históricos e gerais e assim à formulação adequada da teoria linguística (tradução nossa).

A definição de Lightfoot, notadamente direcionada para a questão da origem das mudanças, centraliza o foco de estudo no interior da própria gramática, argumentando ser a própria mudança o teste para se chegar ao limite de possíveis gramáticas. A definição de Labov, por outro lado, e, neste sentido, entenda-se a correlação com a evolução biológica proposta pelo Autor, aparece direcionada fundamentalmente para a questão da propagação, equacionando mesmo origem e propagação. Situando-se, pois, a conceituação de mudança linguística nos dois extremos da origem e da propagação, vê-se claramente quão minimizada pode ser a distância entre os dois quadros teóricos (TARALLO, 1991, p. 13-14).

A aproximação dos dois modelos teóricos em trabalhos investigativos sobre a mudança linguística é outra perspectiva, apontada por Tarallo (1991), e já citada anteriormente. O desenvolvimento do modelo sintático paramétrico abriu espaço, conforme afirmação de Lightfoot (1989), principalmente nos estudos de orientação diacrônica, para a possibilidade de interferência de fatores de natureza não-gramatical na resolução da mudança: “Isto não significa que o modelo procure investigar tais forças; ao contrário, alerta-se tão e simplesmente para um cuidado teórico no sentido de a teoria da gramática não sofrer exigências extremas por parte do teorizador” (TARALLO, 1991, p. 14).

Por outro lado, da mesma maneira, espera-se encontrar no modelo da implementação, de acordo com Tarallo (1991, p. 14), “a possibilidade de a estrutura interna desempenhar um papel mais preponderante na mudança linguística, independente de forças externas, tais como apresentadas na definição de Labov”. Tarallo (1991) defende, então, para que haja a adequação de modelos teóricos, ser necessário reconsiderar a diferenciação entre origem e propagação de mudanças: “Vimos que tal diferenciação, quando e se registrada, equaciona um menor afastamento conceitual e a nível de pressupostos entre modelos teóricos aparentemente incompatíveis” (TARALLO, 1991, p. 19-20).

Por isso mesmo, o desenvolvimento teórico da linguística moderna, conforme afirma Lucchesi (2011), em grande parte, está centralizado no desafio em tentar reunir essas duas dimensões do fenômeno linguístico que, aparentemente, seriam antitéticas e que representam as dimensões psíquico-biológica e sócio-histórica da linguagem humana, sendo que:

No plano formal da estrutura linguística, as contribuições da gramática gerativa apresentam-se como o terreno mais promissor. Entretanto, o grande desafio é abandonar o *axioma da invariância*. A incapacidade da teorização formalista em reconciliar o funcionamento da língua e a variação e a mudança inerentes às condições concretas de uso da língua resultam

de dois grandes equívocos em sua concepção axial da língua: um sistema perfeito que é utilizado somente para transmitir informação referencial (LUCCHESI, 2011, p. 229).

Na concepção gerativista de língua, a língua ou a gramática, não é um sistema perfeito, de acordo com Lucchesi (2011, p. 229), “no sentido em que o são os sistemas das linguagens artificiais”. Essa imperfeição da linguagem humana é que vai torná-la capaz de desempenhar também outras funções sociais que serão tão essenciais quanto a função da transmissão de informações entre os indivíduos:

Uma adequada descrição da estrutura gramatical da língua não pode simplesmente abstrair a variação. Para descrever adequadamente a estrutura invariante da gramática, é necessário definir com precisão onde se situam as suas zonas de variação. E estudos recentes desenvolvidos no âmbito da própria Gramática Gerativa têm revelado que essas zonas de variação se situam precisamente naqueles mecanismos que, na visão chomskiana, seriam os responsáveis pelo caráter imperfeito da linguagem humana: os mecanismos que fazem parte da gramática das línguas humanas, mas que não servem para transmitir qualquer informação referencial (LUCCHESI, 2011, p. 230).

É necessário buscar, então, uma análise que defina precisamente os limites da variação e da invariância na estrutura da gramática e que seja capaz, ainda, para Lucchesi (2011, p. 230), de “integrar a análise gramatical na matriz mais ampla do conjunto complexo de funções que a linguagem humana desempenha no meio social” e que não se restrinja à função comunicativa da linguagem, já que essa seria insustentável, empiricamente falando. Os elementos essenciais para que se possa conceber a integração dessas duas dimensões do fenômeno linguístico pode ser encontrada no debate existente na linguística moderna, desde a década de 1960, que embasa a ideia de que os modelos teóricos da sociolinguística variacionista e da gramática gerativa não são inflexíveis entre si. Isso, segundo Lucchesi (2011), é baseado na utilização de conceitos comuns, que são cruciais para se construir o objeto de estudo em cada um dos dois paradigmas.

Essa afirmação fica evidente em um dos principais trabalhos da Sociolinguística Variacionista, os *Fundamentos Empíricos para uma teoria da Mudança Linguística*, no qual a mudança linguística é concebida racionalmente através da descrição ordenada da diferenciação em uma língua que serve a uma comunidade. Por isso mesmo, “O comando nativo das estruturas heterogêneas não é matéria de multidialectalismo ou de “mero” *desempenho*, mas é parte da *competência* linguística monolíngue” (WEINREICH; LABOV; HERZOG, 2006 [1968],

p. 36). Sendo assim, o fato de os sociolinguistas terem adotado conceitos como competência e desempenho, considerados centrais na reflexão dos gerativistas, e o termo *code-switching*, para indicar a alternância de código, ao discutir a natureza do sistema linguístico, é uma indicação de que o diálogo entre os dois modelos é uma realidade objetiva.

Os modelos teóricos da sociolinguística variacionista e da gramática gerativa estariam reunidos decisivamente num mesmo campo da ciência da linguagem, através da “adoção da noção de estrutura (ou sistema) como um elemento chave em qualquer algoritmo explicativo da linguagem humana” (LUCCHESI, 2011, p. 232). Ao reunir os dois modelos, esses estariam opostos, então, à Análise da Conversação e à Análise do Discurso, correntes da linguística contemporânea, definidas como pós-modernas, conforme enfatiza Lucchesi (2011). A partir desse antagonismo existente no cenário da linguística contemporânea, considerados irreduzíveis entre si, entre a linguística moderna e a linguística pós-moderna, a gramática gerativa e sociolinguística variacionista estariam juntas no campo da modernidade:

Colocamos a Gramática Gerativa e Sociolinguística Variacionista no campo da modernidade, como as duas vertentes que sucedem o estruturalismo linguístico, negando-o, mas mantendo-se no campo de uma ciência da linguagem que pretende uma apreensão global e sistemática do fenômeno linguístico através do conceito de uma estrutura (ou sistema) que possui uma lógica própria; ficando em questão apenas a natureza e o escopo dessa lógica sistêmica da linguagem, bem como a sua localização empírica, como se pode ver no debate que sobretudo a Sociolinguística tem travado com o Gerativismo e a tradição formalista desde o seu texto inaugural (LUCCHESI, 2011, p. 233).

A eleição do idioleto utilizada como uma maneira de isolar o sistema linguístico homogêneo e unitário, como consequência da contradição que foi estabelecida por Saussure (2000 [1916]) entre sistematicidade e variabilidade, é um dos questionamentos primordiais da Sociolinguística em relação à tradição formalista. Conforme afirmam Weinreich, Labov e Herzog (2006 [1968]), a separação entre variação e sistema era defendida tanto por Hermann Paul quanto por Saussure, indo mais além com os seus sucessores que postulavam, cada vez mais, sistematicidade na língua. Isso levou a um maior comprometimento com a visão simplista do idioleto homogêneo, o que vai coincidir absolutamente com o falante-ouvinte ideal de Chomsky (1975 [1965]), que pertence a uma comunidade linguística homogênea. Essa noção serviu de base para o que Chomsky (1988 [1981] e 1994

[1986]) definisse língua-I e língua-E. Essa oposição, como já discutido, estabelece que a primeira fosse interna, intencional e individual e a segunda, objetivada, coletiva e heteróclita. A língua-E seria, então, o objeto de estudo da Sociolinguística (cf. Lucchesi, 2011, p. 234).

As contradições que foram introduzidas por Saussure (2000 [1916]) e que percorrem totalmente a pesquisa linguística do estruturalismo, durante os anos 1930 a 1960, contribuíram para essa oposição entre o gerativismo e a sociolinguística, em relação à localização empírica do objeto de estudo da análise linguística.

O nosso trabalho, assim como diversos estudos recentes, pretende contribuir para que essa incompatibilidade sugerida entre a gramática gerativa e a sociolinguística não impeça uma apreciação e compreensão desses dois modelos teóricos, para que possa haver uma possibilidade de diálogo entre ambos, conforme afirmam Tarallo (1987, 1991), Kroch (1989, 2001), Lucchesi (2011) e Duarte (2016). Sendo assim, apresentaremos resultados, de nossos estudos que incorporam a sociolinguística e a gramática Gerativa, através do modelo da Teoria da Variação e Mudança Linguística, utilizando como componente alguns pressupostos da Teoria de Princípios e Parâmetros.

### 1.3 CONCLUINDO A SEÇÃO

Nesta seção, abordamos o surgimento da Teoria da Gramática, que tem o linguista norte-americano Noam Chomsky como um dos principais nomes dentro das pesquisas gerativistas, desde a publicação de *Syntactic Structures*, em 1957, assumindo o aspecto biológico da linguagem. Vimos, então, que a teoria gerativa afirmava que a linguagem é inata ao ser humano, preocupando-se com a explicação e funcionamento de seu aspecto biológico, ou seja, da competência linguística do falante.

Quanto aos estudos gerativistas propostos por Chomsky, vimos as três fases, de acordo com Boeckx e Hornstein (2007), que são: a *fase combinatória*, a *fase cognitiva* e a *fase minimalista*. A partir daí, passamos a uma explanação das ideias de Chomsky e da teoria gerativa, a respeito dos estudos da língua, começando pela oposição entre *competência* e *desempenho*, a faculdade da linguagem e a gramática universal (GU). Abordamos também os conceitos, estabelecidos por Chomsky, de

língua-E e língua-I, entre outros, de suma importância para o entendimento do gerativismo.

Após discorrermos sobre a Teoria da Gramática, seguimos nosso percurso sobre o estudo da língua, partindo para a análise do comportamento social ou o estudo da fala, de acordo com as teorias variacionistas da língua, denominada de sociolinguística. Apesar de ter surgido muito antes, como vimos, a sociolinguística teve seu marco inicial em 1964, através de uma conferência realizada por 25 pesquisadores em Los Angeles, mas teve o americano William Labov como seu principal difusor. Sendo assim, falamos da relação entre Labov e a sociolinguística, os fundamentos teóricos dessa ciência e os principais problemas da pesquisa sociolinguística.

Ao finalizar esta seção, chegamos à sociolinguística paramétrica, mostrando que não há incompatibilidade entre a junção do modelo teórico de Princípios e Parâmetros (P&P), estabelecido por Chomsky (1988 [1981]) e as pesquisas variacionistas elaboradas acerca das línguas românicas, conforme proposto por Tarallo (1987).

## 2 TEORIA DE PRINCÍPIOS E PARÂMETROS

Nesta seção, abordaremos a Teoria de Princípios e Parâmetros, proposta por Chomsky (1988 [1981]), discutindo, mais especificamente, problemas referentes aos parâmetros e à aquisição da linguagem. Abordaremos um dos parâmetros mais estudados dentro da perspectiva da Teoria de Princípios e Parâmetros, o Parâmetro do Sujeito Nulo, no Português do Brasil.

Em linhas gerais, na Teoria de Princípios e Parâmetros, a GU é composta de dois grupos de informações, como se pode supor, sendo eles os princípios e os parâmetros. Enquanto os princípios dizem respeito às semelhanças compartilhadas entre as línguas, os parâmetros são compostos pelas possíveis diferenças entre as diversas línguas existentes. A Teoria de Princípios e Parâmetros, formulada nos anos 1980, vem se mantendo ativa até os dias atuais, indicando que a GU deve ser compreendida, fundamentalmente, como um conjunto de regularidades gramaticais universais, chamados de princípios, e um conjunto limitado de variações linguísticas possíveis, os parâmetros.

A seguir, abordaremos a mudança de um sistema de regras para um novo modelo, fundamentado na Teoria de Princípios e Parâmetros.

### 2.1 A MUDANÇA DE UM SISTEMA DE REGRAS PARA UM MODELO FUNDAMENTADO EM PRINCÍPIOS E PARÂMETROS

Em seu primeiro estágio, a Teoria de Princípios e Parâmetros, de acordo com Kato (2002), concebe os parâmetros como propriedades associadas aos princípios. Já em seu segundo estágio, os parâmetros passam a ser associados a propriedades do léxico, mais especificamente a propriedades do léxico funcional, havendo, nesse caso, uma evolução de um conceito substantivo para um conceito formal de classes funcionais. A preocupação com a diversidade sintática só passa a se manifestar, de forma explícita, nas pesquisas de linguística gerativa no modelo de Princípios e Parâmetros a partir da década de 1980 (CHOMSKY, 1988 [1981], 1982, 1994 [1986]). Antes desse período, a questão principal era determinar os princípios, sem se atentar para o que possibilitava a sua diversidade.

A partir do momento em que a noção de parâmetros é introduzida, há uma grande expansão, conforme Kato (2002), na produção de trabalhos em linguística comparativa, histórica e na área de psicolinguística, sendo que o seu uso passou a

afetar o próprio conceito de princípio, aumentando ou diminuindo o rol de categorias e funções, refinando a relação léxico/sintaxe e alterando a própria arquitetura da gramática.

Por isso mesmo, a mudança do conceito anterior da GU, de um sistema de regras para um modelo fundamentado em princípios e parâmetros, é considerada um dos pontos mais importantes a serem discutidos:

A argumentação é muito mais complexa, porque a teoria é muito mais simples; baseia-se num número reduzido de princípios gerais, que devem ser suficientes para derivar as consequências de sistemas de regras elaborados e específicos das línguas. Esta mudança notável no carácter do trabalho recente reflecte avanços bastante significativos na direcção da adequação explicativa (CHOMSKY, 1994 [1986], p. 152).

Apesar de ter utilizado diversos exemplos em inglês, muitos estudos de grande importância têm sido produzidos em relação a outras línguas, principalmente em se tratando das línguas românicas. Mesmo afirmando que as ideias propostas e desenvolvidas sejam, em certa medida, questionáveis, Chomsky (1994 [1986]) enfatiza que provavelmente não há dúvidas de que a mudança qualitativa da argumentação demonstrada em diversos trabalhos seja a forma adequada de desenvolvimento. Por isso mesmo:

Em termos de segunda mudança conceptual, já não se considera a GU como um sistema que fornece um formato para os sistemas de regras e uma avaliação métrica. A GU consiste, sim, em vários subsistemas de princípios; tem a estrutura modular que se descobre regularmente na investigação dos sistemas cognitivos (CHOMSKY, 1994 [1986], p. 153).

Sendo assim, a maior parte desses princípios associa-se a parâmetros os quais, segundo Chomsky (1994 [1986]), devem possuir a propriedade de ser fixados por meio de experiências bem simples, uma vez que esta já está disponível para a criança. Logo que os valores dos parâmetros são fixados, o sistema total se torna operativo. Chomsky (1994 [1986]) também enfatiza que a GU é um sistema estruturado profundamente, mas “complicado” apenas parcialmente:

O sistema está associado a um conjunto finito de interruptores, possuindo cada um destes um número finito de posições (talvez duas). Exige-se experiência para regular os interruptores. Quando estes estiverem regulados, o sistema funciona (CHOMSKY, 1994 [1986], p. 153).

Ao se falar em teoria da aquisição, estamos falando de uma concepção diretamente relacionada ao que se entende por língua. Chomsky (1994 [1986]) apresenta duas concepções que distinguem claramente duas possibilidades lógicas de se entender a língua. Sendo assim, há uma língua enquanto objeto externo, representando um conjunto de enunciados e, também, uma língua enquanto algo interno à mente, entendida como conhecimento individual, conforme já citado anteriormente. Para Chomsky (1978 [1965]), a teoria da gramática padrão define a competência do falante através de um conjunto de regras, o que remete ao famoso problema de Platão, que leva ao questionamento do aprendizado de uma língua pelas crianças de forma rápida e, relativamente, homogênea, através de um *input* imperfeito, ou seja, quando não há uma hipótese forte acerca da possibilidade de elas encontrarem um conjunto de regras. O problema de Platão será abordado com maior detalhamento mais à frente, quando tratarmos dos parâmetros e gatilhos no processo de aquisição da linguagem. Ao comentar sobre essa questão, o próprio Chomsky (2012 [1987]) afirmou não haver regras, por isso mesmo não haveria nenhuma necessidade de aprendê-las, conforme veremos mais adiante. Por isso mesmo, o conhecimento sintático de um falante adulto passa a ser visto não mais como um conjunto de regras de boa formação, mas sendo constituído de princípios invariantes e propriedades paramétricas. Os valores dessas propriedades paramétricas seriam definidos de acordo com a língua que serviu de *input* no período da aquisição.

A GU pode ser definida como o conjunto de princípios invariantes responsáveis por reger as línguas naturais e os parâmetros que, por sua vez, podem variar interlinguisticamente, de acordo com as opções de presença [+ ou 1] ou ausência [- ou Ø]. Conforme enfatiza Kato (2002, p. 311): “Os Princípios não são aprendidos, quando muito, maturam”. Quanto aos parâmetros, esses já se encontram também previstos, no entanto, necessitam ter a sua definição especificada de acordo com o um valor [+1 ou -0], o que vai ocorrer de acordo o *input* recebido pela criança da comunidade em que vive.

É importante salientar que a GU permitia um número infinito de línguas, conforme era concebida nos estudos iniciais. No entanto, de acordo com Chomsky (1994 [1986]), a nova concepção de GU só vai permitir um número finito de línguas nucleares, com exceção do léxico, em que há parâmetros em número finito, sendo

que cada um deles possui um número finito de valores, o que seria somente uma mudança qualitativa.

Pode-se, então, pressupor as propriedades paramétricas, de acordo com Chomsky (1994 [1986]), como um número finito e pelo modo como ocorrerá a sua fixação, de forma positiva ou negativa, sendo dependente tão somente de dados positivos. A aquisição, então, não está relacionada a correções ou instruções, já que é um processo de seleção não instrucional. Para que ocorra a transição do estado inicial  $E_0$  ( $S_0$ ) para o estado estável  $E_E$  ( $S_S$  - *Steady State*), é preciso somente regular os interruptores:

Podem existir princípios gerais que determinem o modo como os interruptores são regulados, como, por exemplo, o princípio do subconjunto discutido por Berwick (1982), que estipula que se um parâmetro tiver dois valores, + e -, e o valor - gerar o subconjunto adequado das frases gramaticais geradas pela escolha do valor +, então - é o “valor não marcado”, selecionado na ausência de evidência; esta é uma condição necessária e suficiente para se aprender apenas a partir da evidência positiva, na medida em que os parâmetros sejam independentes. Podem também existir princípios específicos de marcação que relacionem os vários parâmetros, que não necessitam de ser totalmente independentes e podem não o ser (CHOMSKY, 1994 [1986], p. 153).

Uma língua particular, ao ser determinada pela fixação dos valores dos parâmetros, define também a estrutura de cada expressão linguística, que pode ocorrer através de um processo computacional bastante complexo. Voltaremos a falar sobre essa questão mais adiante ao abordarmos o subconjunto de Berwick (1982).

Outra questão enfatizada por Chomsky (1994 [1986]) diz respeito à necessidade de “clarificar” as ideias a respeito da idealização saussuriana-bloomfieldiana de uma comunidade linguística homogênea, afirmando que os sistemas chamados pelo senso comum de “língua” admitem exceções, como a morfologia irregular, as expressões idiomáticas, entre outras. Essas exceções, no entanto, não condizem com a ideia de GU baseada em princípios e parâmetros. E acrescenta:

Suponhamos que distinguimos língua *nuclear* de *periferia*, sendo uma língua nuclear um sistema determinado pela fixação de valores para os parâmetros da GU e a periferia tudo aquilo que é adicionado ao sistema realmente representado na mente/cérebro do falante-ouvinte. Esta discussão é interna à teoria: depende crucialmente da formulação da GU. Ultrapassa a idealização anterior, porque, mesmo sob a assunção de

homogeneidade, pode manter-se uma distinção núcleo-periferia (CHOMSKY, 1994 [1986], p.153-154).

Ao comentar sobre a distinção entre núcleo e periferia, Chomsky (1994 [1986]) afirma haver três noções a respeito dessa:

Núcleo *versus* periferia, interno ao núcleo, interno à periferia. A segunda noção tem a ver com os modos como são fixados os parâmetros na ausência de evidência. Quanto à terceira noção, existem, sem dúvida, regularidades significativas, mesmo quando existe um afastamento relativamente aos princípios nucleares (por exemplo, na morfologia dos verbos irregulares em Inglês). Pode acontecer que as construções periféricas estejam relacionadas com as nucleares de um modo sistemático, digamos que por um enfraquecimento de certas condições da gramática nuclear. O problema da formulação destas noções é empírico, embora não seja simples, e existem muitos tipos de evidência que podem ser relevantes para as determinar. Por exemplo, esperar-se-ia que os fenômenos pertencentes à periferia fossem suportados por evidência específica de “densidade” suficiente, que os mesmos variassem entre as línguas e os dialectos, e por aí adiante (CHOMSKY, 1994 [1986], p. 153-154).

É muito difícil solucionar o problema da elaboração de sistemas de regras, segundo Chomsky (1994 [1986]), que alcancem certo grau de adequação descritiva. A mudança de perspectiva para um sistema de princípios e parâmetros é um problema que deve ser considerado, constituindo, assim, uma instigação empírica ainda mais árdua. Caso determinados fenômenos estejam consignados na periferia, deve-se dar conta deles através de um processo computacional, que seja dedutivo, particularmente, partindo de princípios invariáveis que se associem a parâmetros fixados, o que pode vir a ser complexo:

Para além disto, o sistema da GU é uma estrutura intrincada e altamente restringida; alterações mínimas à caracterização dos princípios e dos conceitos têm consequências vastas e complexas para uma língua particular em análise, bem como para outras línguas (CHOMSKY, 1994 [1986], p.154).

Por isso mesmo, qualquer proposta relacionada a esses princípios e conceitos tem que corresponder a uma enorme variedade de dados bem compreendidos atualmente, além de domínios vastos e em constante expansão. Como se sabe quanto a esses aspectos, há uma grande diferença qualitativa do que havia antes, conforme afirma Chomsky (1994 [1986], p. 155): “Estes problemas e alterações estão bastante claros em abordagens como a de Chomsky e Lasnik

(1977), anterior à mudança de perspectiva para uma teoria com princípios e parâmetros”.

Desse modo, as investigações das línguas, a partir do ponto de vista da gramática gerativa, de acordo com Chomsky (1998), permitiram que se confirmasse o quanto a sua diversidade e a sua complexidade tinham sido subestimadas, mesmo sabendo que tais questões podem ser somente aparências superficiais. Essas conclusões consideradas paradoxais vão levar de forma categórica ao principal problema do estudo da linguagem: “Como podemos mostrar que todas as línguas são variações de um mesmo tema e, simultaneamente, registrar fielmente suas intrincadas propriedades de som e significado, superficialmente diversas?” (CHOMSKY, 1998, p. 24).

Para que uma teoria da linguagem humana seja realmente original, é necessário satisfazer duas condições. A primeira é a adequação descritiva, que tem uma relação direta com a gramática de uma língua particular: “A gramática satisfaz essa condição na medida em que dá uma explicação completa e exata das propriedades da língua, daquilo que o falante da língua sabe” (CHOMSKY, 1998, p. 24). Já a adequação explicativa, por sua vez, mantém uma relação mais direta com teoria geral da linguagem, ou seja, a gramática universal: “Para satisfazer essa condição a gramática universal tem de mostrar que cada língua particular é uma manifestação específica do estado inicial uniforme, dele derivada sob as ‘condições de fronteira’, cujas opções são fixadas pela experiência” (CHOMSKY, 1998, p. 24).

Sendo assim, é possível conseguir uma explicação das propriedades das línguas em um nível maior de profundidade, sendo que à proporção que a GU corresponde à condição de adequação explicativa, ela possibilita uma solução para o que é denominado problema lógico da aquisição da língua, mostrando, a princípio, como esse problema pode ser resolvido e, em seguida, fornecendo uma estrutura que permita o estudo do modo como o processo ocorre.

Há uma tensão que praticamente tem fixado as pautas dessas duas tarefas de pesquisas, já que ao se procurar uma adequação descritiva, para Chomsky (1998), pode-se levar a uma complexidade e uma variedade que seriam sempre maiores em relação ao sistema de regras e, por outro lado, ao se buscar a adequação explicativa, é exigida a invariância da estrutura da língua. A solução encontrada para resolver essa tensão, seria desafiar o pressuposto tradicional que se manteve no início da gramática gerativa, de que:

A língua é um sistema complexo de regras, cada regra sendo específica de línguas particulares e construções gramaticais particulares: regras para formar orações em hindi, sintagmas verbais em bantu, passivas em japonês, e assim por diante. Considerações de adequação explicativa indicam que isso não pode estar correto (CHOMSKY, 1998, p. 24).

As tentativas de encontrar propriedades de sistemas de regras que pudessem ser atribuídas à própria faculdade de linguagem foram uma forma de enfrentar esse problema, de acordo com Chomsky (1998). Todos esses esforços contribuíram para uma nova abordagem da linguagem, divergindo de forma muito mais radical da tradição do que a gramática gerativa anterior, denominada de Princípios e Parâmetros, cujo novo enfoque exclui o conceito de regra e construção gramatical, como já citado anteriormente:

Essa abordagem de “Princípios-e-Parâmetros” como tem sido chamada rejeitou inteiramente o conceito de regra e construção gramatical: não há regras para formar orações relativas em hindi, passivas em japonês, e assim por diante. As construções gramaticais familiares são consideradas artefatos, úteis talvez para a descrição informal, mas sem uma posição dentro da teoria. Elas têm um status parecido com o de “mamífero terrestre” ou “animal caseiro de estimação (CHOMSKY, 1998, p. 25).

As regras são, então, decompostas em princípios gerais da faculdade de linguagem que, por sua vez, de acordo com Chomsky (1998), interagem produzindo as propriedades das expressões. Desse modo, o estado inicial da faculdade de linguagem pode ser imaginado como uma rede de relações fixas conectada a um painel de controle, em que as redes de relações seriam os princípios da linguagem e os controles, as opções a serem determinadas pela experiência:

Quando os controles estão fixados de um modo, temos o bantu; quando estão fixados de outro modo, temos o japonês. Cada língua humana possível é identificada como uma fixação particular dos controles – uma fixação de parâmetros, na terminologia técnica. Se o programa de pesquisa for bem-sucedido, deveríamos ser literalmente capazes de deduzir o bantu a partir de uma certa escolha de fixações, o japonês de outra, e assim por diante, para todas as línguas que os seres humanos podem adquirir (CHOMSKY, 1998, p. 25).

É importante salientar que os controles são fixados baseados em informações muito limitadas de que a criança dispõe, conforme indica Chomsky (1998), o que é exigido pelas condições empíricas de aquisição da língua. Também pequenas mudanças na fixação dos controles podem ocasionar uma enorme variedade

aparente nos dados de saída. Isso porque os efeitos podem proliferar através do sistema. De qualquer forma, para Chomsky (1998), seja como for, qualquer teoria genuína tem de captar essas propriedades gerais da linguagem. Por se tratar de um programa, está longe de ser um produto acabado. Sendo assim, as conclusões obtidas de forma hipotética podem não se manter em sua forma atual. O fato é que dentro do domínio desse programa de pesquisa, tem-se como principal tarefa a descoberta dos princípios e parâmetros.

A questão da aquisição da linguagem será tema da subseção a seguir. Abordaremos também os parâmetros e gatilhos no processo de aprendizagem da língua.

## 2.2 PARÂMETROS E GATILHOS NO PROCESSO DE AQUISIÇÃO DA LINGUAGEM

Ao se tratar da aquisição da linguagem, há diversos estudos e uma grande discussão sobre essa questão. Um desses estudos diz respeito ao trabalho de Lightfoot (1991), o qual afirma que tradicionalmente os linguistas assumem que as crianças não adquirem uma língua tendo como base somente a experiência. Elas também facilitariam o processo de aquisição, tendo uma disposição apropriada para aprender.

Estudos realizados por gramáticos gerativistas, no entanto, mostram “procedimentos de descoberta” e “princípios analógicos” de complexidade crescente que são parte da contribuição da criança nesse processo. Esses estudos também contribuíram para a hipótese de uma teoria seletiva de aquisição da linguagem:

We have sought to ascertain what information must be available to children, independent of any experience with language, in order for the eventual mature linguistic capacities to emerge on exposure to some typical triggering experience. Cutting some corners, we have assumed that this unlearned information is genetically encoded in some fashion, and we have adopted (1) as our explanatory model.

(1) a. trigger (genotype ® phenotype)

b. primary linguistic data (Universal Grammar ® grammar)<sup>8</sup>

---

<sup>8</sup> Temos procurado determinar quais informações devem estar disponíveis para as crianças, independentemente de qualquer experiência com linguagem, a fim de que as capacidades linguísticas maduras possam emergir após a exposição a algumas experiências. Tirando alguns detalhes, adotamos (1) como nosso modelo explicativo.

(1) a. gatilho (genótipo ® fenótipo) b. dados linguísticos primários (Gramática Universal ® gramática) (tradução nossa).

(LIGHTFOOT, 1991, p.1).

O intuito é que aspectos relevantes referentes ao genótipo de uma criança sejam especificados, sendo que, em um estado maduro, conforme Lightfoot (1991), tais crianças são expostas a um tipo de experiência desencadeadora, que vai depender do ambiente linguístico em que ela for criada. A GU contém, então, os aspectos do genótipo que são fundamentais para a formação da linguagem, sendo que a “gramática” faz parte da constituição mental de um indivíduo e é, desse modo, uma propriedade fenotípica, que indica a sua capacidade linguística madura. Sendo assim, os dados linguísticos considerados principais são aqueles a que as crianças estarão expostas e que irão determinar ou “acionar” determinado aspecto de suas gramáticas o que a longo prazo terá algum efeito:

The theory is "selective" in the same sense that current theories of immunology and vision are selective and not "instructive." Under an instructive theory, an outside signal imparts its character to the system that receives it, instructing what is essentially a plastic and modifiable nervous system; under a selective theory, a stimulus may change a system that is already highly structured by identifying and amplifying some component of already available circuitry. Put differently: a selective theory holds that an organism experiences the surrounding environment and selects relevant stimuli according to criteria that are already present internally (LIGHTFOOT, 1991, p. 2).<sup>9</sup>

Ao abordarmos a questão da aquisição da língua, é importante salientar que para a Gramática Gerativa, de acordo com os trabalhos de Chomsky (1978 [1965], 1988 [1981]), a linguística é uma ciência natural, que se ocupa do estudo da faculdade da linguagem, caracterizada como um componente da mente/cérebro, dedicada exclusivamente ao estudo da linguagem humana. Dessa forma, conforme já citado anteriormente, a GU é o estado inicial da faculdade da linguagem ( $S_0$ ), e a gramática do indivíduo adulto, o seu estado final ( $S_s$ ).

Para que a aquisição da linguagem seja possível para o ser humano, os princípios e parâmetros são constituídos pela GU como uma faculdade inata e autônoma, que funciona independente de outros módulos da mente/cérebro. Dentro

---

<sup>9</sup> A teoria é "seletiva" no mesmo sentido em que teorias atuais de imunologia e visão são seletivas e não "instrutiva." Sob uma teoria instrutiva, um sinal externo transmite seu caráter ao sistema que o recebe, instruindo o que é essencialmente um sistema nervoso plástico e modificável; sob uma teoria seletiva, um estímulo pode alterar um sistema que já é altamente estruturado, identificando e ampliando algum componente de circuitos disponíveis. Colocado de maneira diferente: uma teoria seletiva sustenta que um organismo experimenta o ambiente e seleciona estímulos relevantes de acordo com critérios que já estão presentes internamente (tradução nossa).

dessa concepção de Chomsky (1988 [1981]), de que a GU é uma faculdade inata, presente no cérebro humano, Lightfoot (1991) comenta sobre a descrição da formação de anticorpos, citada por Jerne (1967), em que, durante o processo, o antígeno seleciona e amplifica anticorpos já existentes:

Similarly, Hubel and Wiesel showed that particular neurons were preset to react only to specific visual stimuli (for example, to a horizontal line); exposure to a horizontal line entails a radical increase in the number of horizontal-line receptors, and a horizontal line can be said to elicit and select specific responses within the organism. Changeux (1980, 1983) argues along similar lines for a theory of "selective stabilization of synapses" whereby "the genetic program directs the proper interaction between main categories of neurons." "However," he continues, "during development within a given category, several contacts with the same specificity may form," and other elements, which are not selected, may atrophy (1980, p. 193)<sup>10</sup> (LIGHTFOOT, 1991, p.2).

Sendo assim, conforme afirma Lightfoot (1991), o aprendizado está associado à ampliação de certas conexões e à eliminação de outras possibilidades.

Como mencionado previamente, a GU é constituída de princípios, que são extremamente estruturados e restritos, e de determinados parâmetros, que são abertos, através da experiência para serem fixados através de *input*, que demonstra as variações entre as línguas, conforme Chomsky (1988 [1981]). Quando isso acontece, estabelece-se uma gramática denominada de *core grammar*. Os princípios e parâmetros consistem no que denominamos de faculdade da linguagem. Ao comentar sobre essa questão, Lightfoot (1991) reafirma que a GU é composta por princípios e parâmetros definidos em um determinado ambiente linguístico:

So the environment may be said to "select" particular values for the parameters of Universal Grammar. Universal Grammar must be able to support the acquisition of any human grammar, given an appropriate triggering experience. Of course, Universal Grammar need not be seen as homogeneous; it may emerge piecemeal, parts of it being available only at certain stages of a child's development. Grammars are not only attainable under normal childhood conditions; they are also usable for such purposes as speech production and comprehension, they are vulnerable to the kinds

---

<sup>10</sup> Da mesma forma, Hubel e Wiesel mostraram que neurônios particulares estavam prontos para reagir apenas a estímulos visuais específicos (por exemplo, para uma linha horizontal); exposição a uma linha horizontal implica um aumento radical do número de receptores de linha horizontal, e uma linha horizontal pode ser dito para elicitar e selecionar respostas específicas dentro do organismo. Changeux (1980, 1983) argumenta ao longo de linhas semelhantes para uma teoria de "estabilização seletiva de sinapses" em que o programa genético direciona a interação adequada entre as principais categorias de neurônios. "No entanto", continua ele, "durante o desenvolvimento dentro de uma determinada categoria, vários contatos com a mesma especificidade podem formar", e outros elementos, que não são selecionados, podem atrofiar (tradução nossa).

of aphasias that are actually found, and one expects that they will provide part of the basis for understanding the developmental stages that children go through (LIGHTFOOT, 1991, p. 2-3)<sup>11</sup>.

Um ponto importante sobre a aquisição da língua pela criança é a questão abordada por Chomsky (1978 [1965], 1994 [1986]), conhecida como “Problema de Platão”, a qual perpassa por toda a sua trajetória de linguista, desde os seus primeiros trabalhos, sendo o elemento norteador de suas investigações científicas sobre a aquisição da língua. Esse problema, conhecido também como problema lógico de aquisição da linguagem, tem como base o diálogo Mênon, de Platão, em que Sócrates afirma para Mênon que, se estimulado de forma correta, até um escravo analfabeto possui noções de matemática, o que o permitiria realizar alguns cálculos geométricos intuitivamente. Desse modo, para Chomsky (1978 [1965], 1994 [1986]), o problema de Platão nos remete ao fato de que a aquisição da língua materna pela criança ocorre em um período restrito, sem que essa precise se esforçar para que esse processo aconteça, independente da pobreza de estímulo.

O "problema lógico da aquisição da linguagem", que teria fornecido uma grande contribuição para a questão do refinamento empírico referente ao gatilho e aos dados linguísticos primários segundo Lightfoot (1991), levou os gramáticos a pressuporem princípios e parâmetros específicos ao nível da GU. A pobreza do estímulo ou experiência desencadeada pelas crianças, de acordo com Lightfoot (1991), parece não ser suficiente para que sejam determinados todos os aspectos da maturação das capacidades que são atingidos por elas:

It is too poor in three distinct ways: (a) The child's experience is finite, but the capacity eventually attained ranges over an infinite domain and therefore must incorporate some recursive property not demanded by experience. (b) The experience consists partly of degenerate data which have no effect on the emerging capacity [...]. (c) Most important, it fails to provide the data needed to induce many principles and generalizations manifested by the mature capacity (LIGHTFOOT, 1991, p. 3).<sup>12</sup>

---

<sup>11</sup> Assim, pode-se dizer que o ambiente "seleciona" valores particulares para os parâmetros de Gramática Universal. A Gramática Universal deve ser capaz de suportar a aquisição de qualquer gramática humana, dada uma experiência de desencadeamento apropriada. Naturalmente, a Gramática Universal não precisa ser vista como homogênea; pode emergir em partes, apenas disponível em determinados estágios do desenvolvimento de uma criança. Gramáticas não são apenas atingíveis em condições normais de infância; elas também são utilizáveis para propósitos como produção e compreensão de fala, elas são vulneráveis aos tipos de afasia que são realmente encontrados, e espera-se que elas forneçam parte da base para a compreensão do desenvolvimento de estágios pelos quais as crianças passam (tradução nossa).

<sup>12</sup> É muito pobre de três maneiras distintas: (a) a experiência da criança é finita, mas a capacidade eventualmente alcançada varia ao longo de um domínio infinito e, portanto, deve incorporar alguma

Em termos quantitativos e qualitativos, o que a criança ouve em seu convívio com familiares e no grupo social do qual faz parte, é apenas um pequeno reflexo do que é capaz ao utilizar a língua, conforme enfatizam Carvalho e Sousa (2018, p. 10): “Durante o período crítico da aquisição da língua, a criança provavelmente não ouvirá um único exemplo de muitas construções que ela acabará por produzir e processar sem nenhuma dificuldade”. Ainda que a criança ouça expressões que não sejam bem formadas, gramaticalmente falando, ela realizará a exclusão dessas, através dos dados linguísticos primários e, baseado nesses, desenvolverá a gramática de sua língua materna.

Não há informação sistemática para a criança de que não existam determinadas formas, como afirma Lightfoot (1991), ou mesmo que elas sejam “não-gramaticais”, sendo que a evidência irrefutável da não ocorrência de certas formas não está presente na experiência normal da infância. Apesar da não existência de um estímulo para a formação de uma estrutura de dependência, as crianças fazem uso sistematicamente dessa operação. Lightfoot (1991), então, argumenta que tudo indica não haver aprendizado nesse contexto e afirma, além disso, que as operações de movimento ocorridas na gramática são sempre dependentes da estrutura:

If this is invariant and not deducible directly from childhood experience, it is reasonable to suppose that structure dependence is part of what the mind brings to the analysis of experience, not something hypothesized on the basis of evidence. One arrives at this conclusion by noting how the stimulus fails to determine certain aspects of mature grammars, and any such argument clearly depends on assumptions about the triggering experience here (LIGHTFOOT, 1991, p. 4)<sup>13</sup>

Sendo assim, Lightfoot (1991, p. 4) afirma que: “arguments from the poverty of the stimulus reveal not only gross properties of Universal Grammar, such as structure

---

propriedade recursiva não exigida pela experiência. (b) A experiência consiste em parte de degenerar dados que não têm efeito sobre a capacidade emergente [...]. (c) O mais importante, falha em fornecer os dados necessários para induzir muitos princípios e generalizações manifestados pela capacidade madura (tradução nossa).

<sup>13</sup> Se isto é invariável e não dedutível diretamente a partir da experiência da infância, é razoável supor que a dependência da estrutura é parte do que a mente traz para a análise da experiência, não algo hipotetizado com base em evidências. Chega-se a essa conclusão observando como o estímulo falha em determinar certos aspectos das gramáticas maduras, e qualquer argumento desse tipo claramente depende de suposições aqui sobre a experiência de gatilho (tradução nossa).

dependence, but also more fine-grained aspects of actual structures”<sup>14</sup>. Para ilustrar a sua afirmação, Lightfoot (1991) cita ainda Baker (1978) e uma discussão de Hornstein e Lightfoot (1981) e Lightfoot (1982), enfatizando que há muito tempo se concorda que a composição das expressões linguísticas é formada por subunidades que possuem uma estrutura interna e é um consenso que uma gramática não é composta somente por uma lista de expressões, mas compreende um sistema algébrico finito, que, por sua vez, pode originar uma sucessão infinita de expressões.

Quanto à informação da não existência de determinados dados, para Lightfoot (1991), em geral, tais elementos negativos não estariam disponíveis para a criança. Já a ambiguidade, só é percebida mais tarde, já que é caracterizada uma habilidade mais sofisticada que vem a se desenvolver posteriormente e de maneira não uniforme, sendo que a maioria passa despercebida, levando o indivíduo a fazer uso dos significados apropriados que estão disponíveis. Para Lightfoot (1991, p. 6): “To be sure, children come to know these things, and this knowledge is part of the output of the language-acquisition process; but it is not part of the input, not part of the “evidence” for the emerging system, and thus not part of the triggering experience”<sup>15</sup>. O estímulo não é suficiente para definir com totalidade a análise emergente, de acordo com Lightfoot (1991), o que leva as crianças a buscar uma outra base não-experimental.

A ordem linear dos constituintes, por exemplo, é formada como um parâmetro, segundo Lightfoot (1991), o qual é definido na exposição a algum gatilho. Apesar de essa questão estar muito bem estabelecida em relação à GU e no que diz respeito às gramáticas particulares, Lightfoot (1991) questiona a respeito da experiência desencadeadora, ou seja, ao modo como os parâmetros seriam definidos em relação às gramáticas específicas dessas línguas. Há, de acordo com o seu questionamento, uma falha ao se considerar a experiência desencadeadora:

This failure to attend to the triggering experience is remarkable in that linguists have been talking for more than a decade of parameter-setting models of language acquisition, often advertising them as a major

---

<sup>14</sup> Argumentos da pobreza do estímulo revelam não apenas propriedades grosseiras da Gramática Universal, tal como a dependência da estrutura, mas também aspectos mais refinados de estruturas reais (tradução nossa).

<sup>15</sup> Para ter certeza, as crianças chegam a conhecer essas coisas, e esse conhecimento faz parte da saída do processo de aquisição da linguagem; mas não faz parte da entrada, não faz parte da “evidência” para o sistema emergente e, portanto, não faz parte da experiência desencadeadora (tradução nossa).

conceptual change from earlier acquisition models based on evaluation metrics that permitted the "selection" of certain grammars. As parameters are proposed to account for differences between languages, the triggering experience is rarely discussed; where there is discussion, it almost never goes beyond pointing to a class of expressions which are alleged to set the parameter in the appropriate fashion (LIGHTFOOT, 1991, p. 8).<sup>16</sup>

O papel da GU na aquisição da linguagem é o de possibilitar ao aprendiz dispor das possíveis gramáticas no meio em que o sujeito se encontra, sendo que os princípios da GU estão exclusivamente relacionados à aprendizagem das línguas naturais, não se aplicando, assim, a outros domínios cognitivos. Desse modo, a GU busca explicações para a compreensão do modo como os aprendizes da L1, mesmo expostos a um *input* relativamente pobre, conseguem adquirir uma gramática tão complexa. Conforme salienta Grédis (2016, p. 166): “A GU seria, portanto, como uma ponte que liga o input recebido pelo aprendiz à gramática final que ele está adquirindo em sua língua materna”. Importante salientar que diversos estudiosos da obra de Chomsky – Towell e Hawkins (1994), Gregg (1989), Flynn (1996) e White (1987) – afirmam que a GU não possui explicação para tudo, embora ressaltem a sua relevância para os estudos de aquisição da língua e a importância de uma teoria gramatical como parte fundamental dessas pesquisas (cf. GRÉDIS, 2016, p. 165).

Quanto aos dados linguísticos primários que irão desencadear o desenvolvimento de uma gramática infantil, Lightfoot (1991) defende que as crianças não podem ser consideradas “pequenas linguistas”, sendo que, desse modo, elas não constroem sua gramática do mesmo modo que os linguistas constroem suas hipóteses: “For example, the primary data do not include well-organized paradigms or comparable data from other languages”<sup>17</sup> (LIGHTFOOT, 1991, p. 10). Seria em virtude disso que os pais, ao corrigirem determinados aspectos relacionados à fala de seus filhos, fornecendo dados negativos, não indica que essa seja a base geral para o desenvolvimento da linguagem, de acordo com Lightfoot (1991), já que essa correção não acontece com todas as crianças, não sendo, então, indispensável para

---

<sup>16</sup> Esta falha em atender a experiência desencadeadora é notável, já que linguistas têm falado por mais de uma década de modelos de definição de parâmetros de aquisição da linguagem, anunciando-os como uma grande mudança conceitual de modelos de aquisição anteriores baseados em métricas de avaliação que permitia a "seleção" de certas gramáticas. Como os parâmetros são propostos para explicar as diferenças entre línguas, a experiência desencadeadora é raramente discutida; onde há discussão, quase nunca vai além apontando para uma classe de expressões que supostamente definem o parâmetro da maneira apropriada (tradução nossa).

<sup>17</sup> Por exemplo, os dados primários não incluem paradigmas bem-organizados ou dados comparáveis de outras línguas (tradução nossa).

que a linguagem se desenvolva. Além disso, Lightfoot (1991) salienta que, quando essa correção acontece, há uma resistência habitual, comprovada de imediato pelos pais.

McNeill (1966, p. 69 *apud* LIGHTFOOT, 1991, p. 10) apresenta um registro que ilustra essa tão falada resistência:

Child: Nobody don't like me  
 Mother: No, say "nobody likes me."  
 Child: Nobody don't like me  
 (eight repetitions of this dialogue)  
 Mother: No, now listen carefully; say "nobody likes me."  
 Child: Oh, nobody don't likes me<sup>18</sup>.

Esse tipo de correção é feito somente para uma restrita faixa de erros, conforme enfatiza Lightfoot (1991), que se refere, geralmente, a formas morfológicas. Há também argumentos de que mesmo que as crianças não recebam dados negativos diretamente, elas poderiam ter acesso a eles. Sobre essa questão, Chomsky (1988 [1981]) afirma:

If certain structures or rules fail to be exemplified in relatively simple expressions, where they would be expected to be found, then a (possibly marked) option is selected excluding them in the grammar, so that a kind of 'negative evidence' can be available even without corrections, adverse reactions, etc. (CHOMSKY, 1988 [1981], p. 9)<sup>19</sup>.

O gatilho está diretamente relacionado a um grupo aleatório de enunciados, que a criança ouve com frequência, de acordo com Lightfoot (1991), ocorridos em uma estrutura satisfatória de um contexto, consistindo em dados importantes, excluindo assim os dados que não contenham informações significantes, como os de determinadas expressões que não são realizadas. Além do mais, de acordo com Lightfoot (1991, p. 10), "the trigger consists only of simple, unembedded material, and that everything can be learned from structures of 'degree-0 complexity'"<sup>20</sup>.

---

<sup>18</sup> Criança: Ninguém não gosta de mim  
 Mãe: Não, diga "ninguém gosta de mim".  
 Criança: Ninguém não gosta de mim  
 (oito repetições deste diálogo)  
 Mãe: Não, agora ouça atentamente; diga "ninguém gosta de mim".  
 Criança: Oh, ninguém não gosta de mim (tradução nossa).

<sup>19</sup> "Se certas estruturas ou regras deixam de ser exemplificadas em expressões relativamente simples, onde seria de esperar que fossem encontradas, então uma opção (possivelmente marcada) é selecionada excluindo-as na gramática, para que um tipo de 'evidência negativa' possa ser disponível mesmo sem correções, reações adversas, etc." (tradução nossa).

<sup>20</sup> O gatilho consiste apenas em coisas simples, material não embutido, e que tudo pode ser aprendido a partir de estruturas de "grau-0 de complexidade" (tradução nossa).

Por outro lado, ao tratar de como o estudo internalista da linguagem se relaciona com o mundo externo, Chomsky (1998) afirma que tais questões estão inseridas em duas categorias referentes: (i) às relações entre mente e cérebro e (ii) a questões de uso da língua.

O estudo internalista da linguagem tenta descobrir as propriedades do estado inicial da faculdade da linguagem e os estados que este assume sob a influência da experiência. Os estados inicial e atingido são estados do cérebro em primeiro lugar, mas descritos abstratamente, não em termos de células, mas em termos de propriedades que os mecanismos do cérebro têm de satisfazer de algum modo (CHOMSKY, 1998, p. 26-27).

Essas duas questões básicas da linguagem são resumidas por Chomsky (1998) ao afirmar que, de um lado, a questão internalista analisa que tipo de sistema é a linguagem e que, por outro lado, é a questão externalista que analisa o modo como a linguagem se relaciona com as outras partes da mente e com o mundo externo, o que inclui problemas de unificação e de uso da língua. Ao examinar mais detalhadamente a questão internalista, Chomsky (1998) salienta que o estudo da linguagem, desenvolveu novos trabalhos acerca do tema, a partir dos anos 1950, o que, além de retomar discussões prévias, reformulou questões e indagações tradicionais acerca de muitos tópicos, o que incluiu a língua e seu uso bem como a relevância desses temas para o estudo da mente humana. Os limites de compreensão e as barreiras conceituais que impediam as tentativas anteriores de se explorar esses temas foram superados em parte, o que permitiu que se prosseguisse mais proveitosamente:

O programa de pesquisa que se desenvolveu conduziu finalmente a um quadro da linguagem que representa uma considerável divergência da longa e rica tradição: a abordagem de Princípios-e-Parâmetros, que se baseia na idéia de que o estado inicial da faculdade de linguagem consiste em princípios invariantes e em um leque finito de escolhas quanto ao funcionamento do sistema inteiro. Uma língua particular é determinada fazendo-se escolhas de um modo específico. Temos aí, pelo menos, as linhas gerais de uma verdadeira teoria da linguagem, que talvez seja capaz de satisfazer as condições de adequação descritiva e explicativa e de abordar o problema lógico da aquisição da língua de modo construtivo (CHOMSKY, 1998, p. 39-40).

A partir do momento que esse cenário ganhou forma, todos os esforços dos estudos e pesquisas se concentraram em revelar e explicitar os princípios e parâmetros, sendo que essas investigações se estenderam muito rapidamente, de acordo com Chomsky (1998), tanto de maneira mais profunda em línguas

individuais, quanto em uma esfera maior, onde ideias semelhantes foram aplicadas a línguas de uma série de categorias mais amplas. A mente/cérebro do ser humano pode ser considerada como o mais complexo objeto no universo, sendo que ainda mal iniciamos a compreensão dos modos como ela se constitui e funciona:

Dentro dela, a linguagem parece ocupar um lugar central, e, pelo menos na superfície, a variedade e a complexidade são desencorajadoras. No entanto, tem havido muito progresso – o bastante para que pareça razoável considerar algumas questões de maior alcance sobre a configuração geral da linguagem, em particular, questões sobre a otimidade da configuração geral (CHOMSKY, 1998, p. 40).

Sendo assim, Chomsky (1998) nos instiga a questionar até que ponto a linguagem é bem configurada e poderia se assemelhar a algo construído por um engenheiro extremamente competente, de acordo com certas especificações de uma configuração geral. Para isso, seria necessário estudar melhor essa questão e explicitar tais especificações:

Algumas são internas e gerais, tendo a ver com a naturalidade conceitual e a simplicidade, noções que dificilmente são límpidas, mas que podem ser avivadas de vários modos. Outras são externas e específicas, tendo a ver com as condições impostas pelos sistemas da mente/cérebro com que a faculdade da linguagem interage (CHOMSKY, 1998, p. 40).

Uma provável resposta a essa questão seria, então, de acordo com Chomsky (1998), a configuração quase perfeita da linguagem em relação à satisfação das condições externas. E se essa conclusão é verdadeira, ela também surpreende por diversas razões, já que a forma como as línguas têm sido vistas, supostamente como objetos altamente complexos e defectivos, estudá-las sob uma perspectiva teórica rigorosa nem valeria a pena, por isso: “Elas exigem reforma ou sistematização, ou substituição por algo bem diferente, se têm de servir a algum propósito, além dos confusos e intrincados assuntos do cotidiano” (CHOMSKY, 1998, p. 41). Sendo assim, tentativas tradicionais de inventar uma língua universal perfeita ou de se recuperar uma língua primitiva original foram inspiradas por essa ideia norteadora.

A partir da próxima subseção, ampliaremos as nossas discussões a respeito do Parâmetro do Sujeito Nulo, devido à sua relevância dentro da Teoria Gerativa, e por exemplificar mais detalhadamente esse processo. A existência da posição de sujeito nas orações nas línguas humanas, que pode ser preenchido foneticamente

ou não, dividirá as línguas em *pro-drop* e *não-pro-drop*, ou seja, línguas de sujeito nulo e línguas não sujeito nulo, conforme veremos a seguir, ao abordarmos o Parâmetro do Sujeito Nulo.

### 2.3 O PARÂMETRO DO SUJEITO NULO

Um dos exemplos mais citados, em se tratando da Teoria de Princípios e Parâmetros, é o Parâmetro do Sujeito Nulo. As línguas humanas são compostas por sujeitos sintáticos e seus respectivos predicados, de acordo com um dos Princípios da GU. Esse é um dos princípios rígidos da GU: a existência, nas línguas humanas, da posição de sujeito nas orações, o que não indica que essa posição tenha que ser preenchida obrigatoriamente por um NP com conteúdo fonético. O natural, então, é que as frases de todas as línguas naturais sejam compostas por sujeitos, o que é observado quando se analisam dados de milhares de línguas do mundo, já que em todas elas as frases são construídas por meio da articulação entre um sujeito e um predicado. Entretanto, em algumas línguas, poderá ocorrer a omissão do sujeito na frase, originando o sujeito nulo.

Línguas como o português, o italiano e o espanhol, designadas *pro-drop*, permitem que essa posição fique vazia, conforme Chomsky (1988 [1981]), sem necessariamente estar preenchida com o sujeito fonético. Por outro lado, línguas como o francês e o inglês, que são designadas *não pro-drop*, não permitem que a posição de sujeito não seja preenchida foneticamente, o que é afirmado e exemplificado também por Lightfoot (1991) e Roberts (2007), mais à frente. Como esse fenômeno é variável, essa possibilidade configura um dos Parâmetros da GU, o Parâmetro do Sujeito Nulo. Desse modo, enquanto parâmetro, o Sujeito Nulo possui em sua essência uma variabilidade binária entre as línguas, de acordo com Chomsky (1988 [1981]), isto é, algumas línguas irão marcá-lo como positivo e outras, como negativo. Quando esse parâmetro é marcado como positivo numa língua, dizemos que ela é [+ sujeito nulo] e quando é marcado como negativo, a língua é considerada [- sujeito nulo].

As propriedades relacionadas ao parâmetro *pro-drop* são o fator mais relevante, segundo Chomsky (1988 [1981]), quando se fala na Teoria de Princípios e Parâmetros. Ao comentar sobre as línguas designadas *pro-drop*, como o italiano, o

espanhol e o português, o autor salienta que há uma tendência a encontrar, entre outras, as seguintes propriedades sintáticas:

- (i) sujeito nulo;
- (ii) inversão livre do sujeito em orações principais;
- (iii) movimento wh "longo" do sujeito;
- (iv) pronomes lembretes nulos na oração encaixada;
- (v) aparente violação do filtro "that-*t*" (CHOMSKY, 1988 [1981], p. 240).

As línguas não *pro-drop*, como o francês e o inglês, por exemplo, de acordo com Chomsky (1988 [1981]), não possuem todas essas propriedades, especificamente. Para as propriedades elencadas acima, são apresentados os seguintes exemplos do italiano:

- (i) ho trovato il libro  
("Achei o livro")
- (ii) ha mangiato Giovanni  
("Giovanni comeu")
- (iii) l'uomo [che mi domando [chi abbia visto]]  
(com a interpretação: "o homem x tal que eu me pergunto quem x viu")
- (iv) ecco la ragazza [che mi domando [chi crede [che possa VP]]]  
("aqui está a menina que eu me pergunto quem acha que pode VP")
- (v) chi credi [che partirà]  
("quem você acha [(que) partirá]") (CHOMSKY, 1988 [1981], p. 240).

Por isso mesmo, supõe-se que há um parâmetro da GU responsável por distinguir, por exemplo, as línguas do tipo italiano, espanhol e português, de línguas como o francês e o inglês: o parâmetro *pro-drop*. Ainda segundo Chomsky (1988 [1981]), a partir do momento em que esse parâmetro é definido, acontece também o agrupamento das propriedades citadas anteriormente. Durante o aprendizado da língua, o indivíduo busca evidências para a fixação do parâmetro ao conhecer as outras propriedades da língua que seguem o valor escolhido:

To begin with, let us assume - following Taraldsen (1978b) - that the parameter involves the inflectional element INFL, or more precisely, the agreement element AGR (= PRO) that is the crucial component of INFL with respect to government and binding. The intuitive idea is that where there - is overt agreement, the subject can be dropped, since the deletion is recoverable. In Italian-type languages, with a richer inflectional system, the element AGR permits subject-drop while in French-type languages it does not. A language might have a mixed system, permitting subject drop in some

constructions but not in others, a property that we might expect to find varying as inflection is or is not overt; Taraldsen gives examples from Irish; Hebrew is another case. The correlation with overt inflection need not be exact. We expect at most a tendency in this direction. The idea is, then, that there is some abstract property of AGR, correlated more or less with overt morphology, that distinguishes pro-drop from non-pro-drop languages, from which the clustering of properties in (I) follows (CHOMSKY, 1988 [1981], p. 241)<sup>21</sup>.

A convergência dessas propriedades sintáticas nas línguas *pro-drop*, como indica Chomsky (1988 [1981]), pode ter relação com a capacidade que o elemento *Agr* tem de reger satisfatoriamente a posição de sujeito nesse tipo de línguas, sendo que, em línguas *não-pro-drop*, o *Agr* não estaria habilitado para tal. Sendo assim, a morfologia “rica” das línguas *pro-drop* seria uma das principais causas do licenciamento do sujeito nulo, conforme Chomsky (1988 [1981], p. 241):

For the moment, let us simply say that in Italian-type languages AGR is able to control the subject in some sense, whereas in French-type languages it is not. We continue to assume, then, that all of these languages have the basic core structure (3):

(3)  $S \rightarrow NP\ INFL\ VP$ , where  $INFL = [[\pm Tense], (AGR)]$   
 (AGR = PRO)<sup>22</sup>

Essa suposição, então, produzirá de imediato as propriedades citadas acima, de acordo com Chomsky (1988 [1981]), principalmente no que se refere às propriedades de (i) a (iv), as quais ele se refere para enfatizar que, em um idioma *pro-drop*, o elemento de concordância *Agr* é responsável pelo controle do sujeito nulo:

<sup>21</sup> Para começar, vamos assumir - seguindo Taraldsen (1978b) - que o parâmetro envolve o elemento flexional INFL, ou mais precisamente, o elemento de concordância AGR (= PRO), que é o componente crucial do INFL com relação à concordância e à vinculação. A ideia intuitiva é que, quando houver uma concordância manifesta, o sujeito pode ser descartado, pois a exclusão é recuperável. Nos idiomas do tipo italiano, com um sistema flexional mais rico, o elemento AGR permite a queda do sujeito, enquanto nos idiomas do francês não. Um idioma pode ter um sistema misto, permitindo a queda de sujeito em algumas construções, mas não em outras, uma propriedade que podemos esperar encontrar variando conforme a flexão seja ou não evidente; Taraldsen dá exemplos do irlandês; o hebraico é outro caso. A correlação com a flexão mais precisa não é exata. Esperamos, no máximo, uma tendência nessa direção. A ideia é, então, que existe alguma propriedade abstrata do AGR, correlacionada mais ou menos com a morfologia aberta, que distingue idiomas *pro-drop* de *não-pro-drop*, a partir dos quais o agrupamento de propriedades em (I) segue (tradução nossa).

<sup>22</sup> Por enquanto, digamos simplesmente que, nas línguas do tipo italiano, a AGR é capaz de controlar o sujeito em algum sentido, enquanto nas línguas do tipo francês não é. Continuamos assumindo, então, que todas essas línguas têm o núcleo de estrutura básica (3): (3)  $S \rightarrow NP\ INFL\ VP$ , em que  $INFL = [[\pm tempo], (AGR)]$  (AGR = PRO) (tradução nossa).

In a pro-drop language, the subject [ $NP^e$ ] is locally controlled by AGR in tensed sentences. Therefore the subject may be missing (i.e., may be [ $NP^e$ ]; property (i)); free inversion leaving the trace [ $NP^e$ ] is permitted (property (ii)); long-*wh*-movement is a free option (property (iii)); and an empty resumptive pronoun may appear in an embedded clause (property (iv)). In each case, the element [ $NP^e$ ] in subject position is locally controlled by AGR. In a non-pro-drop language, none of these options exists, since [ $NP^e$ ] will not be locally controlled (CHOMSKY, 1988 [1981], p. 242)<sup>23</sup>.

Um dos elementos do grupo de características relacionadas ao parâmetro *pro-drop* é que as línguas *pro-drop* permitem a inversão com bastante facilidade. Por isso mesmo, Chomsky (1988 [1981]) sugere que esse grupo de propriedades, que são características dos idiomas *pro-drop* citadas acima, seja repensado. Haveria fortes evidências de que os itens (i) e (ii) sejam as propriedades consideradas essenciais em relação às línguas *pro-drop*, ou seja, o sujeito nulo e a inversão livre do sujeito, respectivamente.

Nesse sentido, Chomsky (1988 [1981]), também enfatiza que:

Assuming so, it follows that in the pro-drop languages the subject may fail to be governed by AGR. We still follow Taraldsen's basic intuition that a property of AGR determines the pro-drop parameter, but now inverting the logic, so that AGR in the pro-drop languages is a "weaker" rather than a "stronger" governor for the subject position. Pursuing the intuition that the pro-drop parameter is related to overt inflection, we might say that in the pro-drop languages the element AGR is more closely connected with the verbal element with which it is morphologically manifested, and thus need not govern the subject position (CHOMSKY, 1988 [1981], p. 256)<sup>24</sup>.

A principal circunstância que motiva a hipótese do Parâmetro do Sujeito Nulo, retomando o que foi dito por Chomsky (1988 [1981]), é que enquanto algumas línguas permitem que o sujeito pronominal, definido, referencial, não seja expresso, em outras, isso não seria possível. Como exemplo, temos o caso do italiano, do

---

<sup>23</sup> Em uma língua *pro-drop*, o sujeito [ $NP^e$ ] é controlado localmente pelo AGR em sentenças de tempo. Portanto, o sujeito pode estar ausente (ou seja, pode ser [ $NP^e$ ]; propriedade (i)); é permitida inversão livre deixando o traço [ $NP^e$ ] (propriedade (ii)); o movimento *wh* longo é uma opção livre (propriedade (iii)); e um pronome lembrete nulo pode aparecer em uma oração encaixada (propriedade (iv)). Em cada caso, o elemento [ $NP^e$ ] na posição de sujeito é controlado localmente pela AGR. Em uma língua não-*pro-drop*, nenhuma dessas opções existe, pois [ $NP^e$ ] não será controlado localmente (tradução nossa).

<sup>24</sup> Partindo do pressuposto, segue-se que nas línguas *pro-drop* o sujeito pode falhar em ser controlado pela AGR. Ainda seguimos a intuição básica de Taraldsen de que uma propriedade da AGR determina o parâmetro *pro-drop*, mas agora invertendo a lógica, de modo que a AGR nas línguas *pro-drop* é um controlador "mais fraco" do que um controlador "mais forte" para a posição do sujeito. Prosseguindo a intuição de que o parâmetro *pro-drop* está relacionado à flexão aberta, poderíamos dizer que, nas línguas *pro-drop*, o elemento AGR está mais intimamente conectado ao elemento verbal com o qual ele se manifesta morfologicamente e, portanto, não precisa controlar a posição do sujeito (tradução nossa).

espanhol e do grego, que são línguas de sujeito nulo ou *pro-drop* (*Parla italiano/Habla español/Mila ellinika*), e do inglês e do francês, que não são línguas de sujeito nulo ou *não-pro-drop* (*\*Speaks English/ \*Parle Français*) (Cf. Roberts 2007, p. 25).

Nesse caso, conforme Lightfoot (1991), também retomando o que foi afirmado acima por Chomsky (1988 [1981]) e Roberts (2007), ocorre o que é denominado parâmetro do sujeito nulo, em que frases declarativas finitas ocorrem com um sujeito foneticamente nulo em línguas como o italiano e o espanhol, por exemplo, mas, por outro lado, não ocorrem em línguas como o inglês e o francês, dentre outras, conforme exemplos a seguir, retirados de Lightfoot (1991, p. 11):

- (12) a. ho trovato il libro
- b. chi credi che partirà
- c. \*found the book
- d. \*who do you think that will leave?<sup>25</sup>

Sobre essa questão, independente da forma do parâmetro que venha a permitir essa variação, de acordo com uma sugestão de Chomsky (1978 [1965]), caso a criança falante do inglês escolha a configuração errada, o fato de não serem capazes de ouvir frases como (12c) possivelmente será tomado como uma indicação indireta da agramaticalidade dessas sentenças e que, por algum motivo, não ocorrem.

Do mesmo modo, temos o português como língua [+ sujeito nulo], enquanto o inglês é exemplo de língua [- sujeito nulo]. Em português, uma frase que apresenta o sujeito pleno (explícito) ou a que apresenta o sujeito nulo são consideradas igualmente gramaticais, sendo totalmente aceitáveis. Por outro lado, no inglês, apenas é gramatical a frase que apresenta o sujeito pleno, sendo negativo o seu parâmetro do sujeito nulo.

Por isso mesmo, de acordo com Roberts (2007, p. 25), retomando o que foi afirmado por Chomsky (1988 [1981]) e Lightfoot (1991), sobre línguas de sujeito nulo e línguas não-sujeito-nulo:

- For now I will present parameters in the form of yes–no questions, and so the null-subject parameter may be informally stated as follows:
- A. Does every White clause require an overt subject?

---

<sup>25</sup> Encontrou o livro / Quem você pensa que vai sair? (tradução nossa).

YES: non-null-subject languages (French, English ...).  
 NO: null-subject languages (Italian, Spanish, Greek, Japanese, Navajo ...)<sup>26</sup>

Desse modo, Roberts (2007) aborda os tópicos acima levando em conta como o indivíduo durante o processo de aquisição da linguagem estaria exposto aos dados linguísticos primários dentro do processo de fixação de valores dos parâmetros. Seria possível formular uma apresentação do parâmetro tendo como base o item A, o que equivaleria a afirmar, segundo Roberts (2007, p. 25), que “Toda oração finita tem um sujeito evidente”<sup>27</sup> e considerar essa afirmação como verdadeira em relação ao valor do parâmetro. As duas afirmações, de acordo com o autor, indicam prontamente a natureza binária dos parâmetros. Ainda, segundo Roberts (2007, p. 25-26):

The null-subject parameter refers, as stated above, to finite, discourse-neutral clauses, and involves the interpretation of the null subject as a definite, referential pronoun. Many non-null-subject languages, including English, allow null subjects under other conditions. For example, both English and French extensively allow or require the subject of non-finite clauses to be null: <sup>28</sup>

- (11) a. [ (Him) smoking ] bothers me.  
 b. John expects [ (Mary) to leave soon ].  
 c. Jean a essayé [ de . . . partir ].  
     John has tried [ . . . to leave ].

Em função disso, em exemplos como os citados anteriormente, de acordo com Roberts (2007), os sujeitos apresentam propriedades diferentes dos sujeitos nulos dos exemplos expostos previamente (do italiano, de espanhol e do grego). O que ocorre no caso dos exemplos em 11b e 11c é que o sujeito vazio do infinitivo tem que ser correferente com a oração principal, ou seja, é controlado pelo sujeito.

---

<sup>26</sup> Por enquanto, apresentarei parâmetros sob a forma de *yes-no questions*, e assim o parâmetro de sujeito nulo pode ser informalmente declarado da seguinte forma:

- A. Toda oração finita requer um sujeito evidente?

SIM: línguas de não-sujeito-nulo (francês, inglês ...)

NÃO: línguas de sujeito nulo (italiano, espanhol, grego, japonês, navajo...) (tradução nossa).

<sup>27</sup> tradução nossa.

<sup>28</sup> O parâmetro nulo refere-se, como dito acima, a orações finitas e neutras do discurso, e envolve a interpretação do sujeito nulo como um pronome referencial definido. Muitas línguas não-sujeito-nulo, incluindo o inglês, permitem sujeitos nulos sob outras condições. Por exemplo, tanto o inglês quanto o francês permitem ou exigem que sujeito de orações não finitas seja nulo: (tradução nossa).

Já no caso do exemplo citado em 11a ele pode ser arbitrário. Isso indica que esses exemplos foram analisados diferentemente dos citados. Segundo Roberts (2007), a maioria dos estudos apresenta apenas as análises técnicas padrão das construções aludidas inicialmente. Tanto o inglês quanto o francês podem permitir sujeitos nulos em situações ou registros especiais de discurso, conforme Roberts (2007, p. 26):

Haegeman (2000: 130) gives the following examples from what she calls 'written abbreviated registers' ('written registers in which pressures of economy seem to overrule the "core" grammar' (132), including diaries, short notes and some kinds of colloquial speech):<sup>29</sup>

- (12) a. . . . cried yesterday morning.  
(Plath 1983: 288)  
b. Elle est alsacienne. . . . Paraît intelligente.  
She is Alsatian. . . . Seems intelligent.  
(Léautaud 1989: 48)

Sendo assim, esses tipos de sujeitos nulos, conforme afirma Roberts (2007), além de serem restritos a determinados tipos de discurso e/ou registro, também possuem propriedades específicas que os distinguem daqueles sujeitos nulos canônicos vistos inicialmente, como é o caso do espanhol, italiano, grego etc. Por outro lado, há línguas como o alemão que permitem sujeito nulo expletivo não-referencial, não permitindo sujeito nulo referencial, como afirma Roberts (2007, p. 26):

In (13a) the expletive pronoun *es* is obligatorily 'dropped', while in (13b) the same pronoun in the same syntactic position, only now with a referential interpretation, cannot be dropped (examples from Cardinaletti (1990: 5-6)):<sup>30</sup>

- (13) a. Gestern wurde (\*es) getanzt.  
yesterday was (it) danced  
'Yesterday there was dancing.'  
b. Gestern war \*(es) geschlossen.  
yesterday was (it) closed  
'Yesterday it was closed.'

<sup>29</sup> Haegeman (2000, p. 130) dá os seguintes exemplos do que ela chama de 'registros abreviados escritos' ('registros escritos em que pressões da economia parecem anular a gramática "núcleo" (132), incluindo diários, notas curtas e alguns tipos de discurso coloquial): (tradução nossa).

<sup>30</sup> Em (13a) o pronome expletivo é obrigatoriamente 'descartado', enquanto em (13b) o mesmo pronome na mesma posição sintática, só que agora com uma interpretação referencial, não pode ser descartado (exemplos de Cardinaletti (1990: 5-6) (tradução nossa)).

Por possuírem essas restrições em relação aos sujeitos nulos, essas línguas, de um modo geral, segundo Roberts (2007), no entanto, não são consideradas totalmente nulas ou línguas nulas “completas”. Isso foi identificado por Rizzi (1982) que denominou essa situação como sendo referente a parâmetros relacionados, no entanto, autônomos, já que permitiriam situações distintas, a depender da língua. No caso do inglês, por exemplo, um sujeito nulo não referencial não é permitido, sendo que o mesmo ocorrerá com um sujeito nulo com um pronome referencial. Já no italiano, ambos os casos são permitidos, o que não aconteceria com o alemão, que permite o primeiro caso, mas não admite o segundo.

Rizzi (1982) chamou as línguas como o alemão de *semi-pro-drop*, já que licencia expletivos nulos, embora não admita sujeitos referenciais nulos. Dessa forma, Rizzi (1982) distingue as línguas de sujeito nulo das línguas de sujeito não nulo, apontando de forma discreta uma distinção entre línguas *pro-drop*, que permitiriam todo tipo de sujeito nulo, tanto referencial quanto não referencial, como é o caso do italiano, e aquelas que permitem somente sujeitos nulos não referenciais, ou seja, sujeitos nulos não temáticos, citando como exemplo variedades do holandês.

Em línguas de sujeito nulo, o pronome sujeito preenchido é permitido em situações que são interpretadas como enfáticas, como nos exemplos a seguir do italiano, do espanhol e do grego,<sup>31</sup> retirados de Roberts (2007, p. 27):

- (14) a. Lui parla italiano. (Italian)  
 HE speaks Italian  
 b. Él habla español. (Spanish)  
 HE speaks Spanish  
 c. Aftos mila ellinika. (Greek)  
 HE speaks Greek

Tomando como base o parâmetro do sujeito nulo, isso demanda a noção de que a exata formulação do parâmetro não vai se referir necessariamente na propriedade concreta relacionada à omissão ou não do sujeito, mas em uma propriedade mais abstrata da GU. Essa propriedade é que tem o efeito de possibilitar ou não a omissão do sujeito. A partir dessa noção de parâmetro, é

---

<sup>31</sup> Nesses exemplos, a tradução para o inglês traz o pronome em letras maiúsculas para dar um destaque.

possível deduzir, conforme afirma Raposo (1992, p.62) que “a omissão do sujeito é apenas uma das propriedades concretas que se encontram associadas à fixação do parâmetro na posição *pro-drop*, o que é confirmado pela análise linguística”.

Em detrimento do que foi exposto, o modelo de Princípios e Parâmetros e, mais ainda, a noção de parâmetros, em particular, tem propiciado associar estudos gramaticais considerados até pouco tempo inflexíveis entre si, como o que estamos propondo neste trabalho, ao associarmos a Teoria Gerativa com a Sociolinguística Variacionista de orientação laboviana. A possibilidade desse diálogo, proporcionando uma interdisciplinaridade, apesar de ainda estar no começo, possibilitará o desenvolvimento e a modificação, que serão consideráveis a uma visão específica da Gramática Universal, proposta na Teoria da Regência e Ligação e, principalmente, em relação à noção de parâmetro.

Abordaremos a seguir, o Princípio do Subconjunto para ilustramos melhor a questão do acesso de dados negativos durante o processo de aquisição da língua.

### 2.3.1 O Princípio do Subconjunto e o acesso a dados negativos durante o processo da aprendizagem

Retomando a questão da hipótese, no Parâmetro do Sujeito Nulo, da ausência de informação negativa durante o processo de aprendizagem e de se o parâmetro possui uma posição neutra ou um valor primitivo não marcado, além de quais as implicações que cada uma delas teria em relação a este, Chomsky (1988 [1981]) considera que se no estado inicial  $S_0$  o parâmetro for neutro, em relação a qualquer um dos valores (“+” ou “-”), isso não ofereceria nenhum problema para uma criança que aprende português fixar o parâmetro no valor *pro-drop*. Desse modo, a presença de orações sem a realização do sujeito fonético nos dados iniciais possibilitará a decisão da criança de que a sua língua é realmente *pro-drop*, ligando esse valor ao fenômeno. Nesse caso, a criança fixa o parâmetro baseando-se em uma informação positiva, ou seja, a não realização do sujeito fonético nas orações no meio linguístico em que convive. O mesmo, no entanto, não ocorreria com uma criança que estaria aprendendo o Inglês, por exemplo, o que seria, de certo modo, uma tarefa mais complicada, já que orações sem a realização do sujeito fonético não existem nesta língua, sendo, portanto, agramaticais. No entanto, por se tratar de uma posição negativa, a criança não tem acesso a essa conclusão.

Em certos casos, mesmo que uma determinada estrutura não ocorra nos dados linguísticos primários, através de uma informação negativa indireta, isso é suficiente para que seja fixado o valor de um parâmetro sem que seja incluído nesse processo correções, ou seja, informação negativa direta:

There is good reason to believe that direct negative evidence is not necessary for language acquisition, but indirect negative evidence may be relevant - We would expect the order of appearance of structures in language acquisition to reflect the structure of markedness in some respects, but there are many complicating factors: e.g., processes of maturation may be such as to permit certain unmarked structures to be manifested only relatively late in language acquisition, frequency effects may intervene, etc. (CHOMSKY, 1988 [1981], p. 9).<sup>32</sup>

Sendo assim, a criança que aprende o inglês possivelmente fixará o valor do parâmetro na posição *não-pro-drop*. Isso se dará não porque a criança teve acesso à informação de que orações sem o preenchimento do sujeito fonético não são gramaticais, mas pelo simples fato de nunca ter ouvido orações desse tipo, o que seria suficiente para que ela fixe o parâmetro nessa posição. Por outro lado, quando se trata de um parâmetro com um valor não marcado no estado inicial  $S_0$ , a informação negativa, incluindo aí a negativa indireta, não desempenhará nenhum papel na aquisição da língua, quando o valor não marcado for a posição *não-pro-drop*. Se no processo de aquisição a criança possui o valor do parâmetro do sujeito nulo fixado na posição *não-pro-drop* e cresce numa comunidade em que se fala inglês, ela não tocará no parâmetro, já que no meio linguístico em que convive não há dados positivos para contradizer a posição predeterminada do parâmetro. Contudo, se a criança cresce em uma comunidade falante do português, tendo acesso à frase em que ocorre a não realização fonética do sujeito, ou seja, a uma informação *positiva*, contradizendo a fixação original do parâmetro, ela irá colocar o “interruptor” na posição oposta:

Neste modelo, a informação negativa indirecta não desempenha nenhum papel na aquisição. Em resumo, se adoptarmos o modelo da teoria dos parâmetros em que estes são inicialmente neutros quanto à sua ligação

---

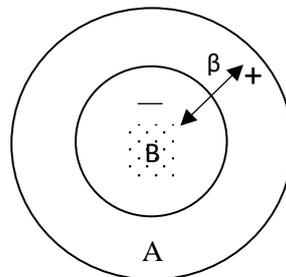
<sup>32</sup> Há boas razões para acreditar que a evidência negativa direta não é necessária para a aquisição da linguagem, mas a evidência negativa indireta pode ser relevante - Esperaríamos que a ordem de aparecimento das estruturas na aquisição da linguagem refletisse a estrutura da marcação em alguns aspectos, mas há muitos fatores complicadores: por exemplo, os processos de maturação podem ser tais que permitam que certas estruturas não marcadas se manifestem apenas relativamente tarde na aquisição da linguagem, efeitos de frequência podem intervir, etc.

final, temos de permitir que a informação negativa indirecta seja relevante na aquisição do sistema linguístico final. A manutenção da versão forte da hipótese sobre a não pertinência da informação negativa exige que adoptemos um modelo específico da teoria dos parâmetros, em que estes possuem um valor inicial não marcado e que apenas é alterado no decurso da aquisição se houver evidência positiva em contrário nos dados primários (RAPOSO, 1992, p. 59).

Berwick (1982) apresentou uma proposta de formalização parcial da hipótese, de acordo com a teoria dos conjuntos, em que afirma não haver relevância dos dados negativos na seleção do valor dos parâmetros. A formalização apresentada por Raposo (1992), em que ele explica o Princípio do Subconjunto, é uma adaptação de Chomsky (1986, p. 146): “Se um parâmetro tem dois valores, “+” e “-”, e o valor “-” determina uma linguagem que é um subconjunto da linguagem determinada pelo valor “+”, então o valor “-” é o valor não-marcado, selecionado por UG na ausência de qualquer evidência linguística”.

Através da ilustração abaixo, reproduzida de Raposo (1992, p. 60), do diagrama de Venn, em que a língua A inclui todas as propriedades da língua B, é explicado o Princípio do Subconjunto:

**Figura 1-** Princípio do subconjunto



**Fonte:** Raposo (1992, p. 60)

O Princípio do Subconjunto indica que “o parâmetro  $\beta$  tem o valor não marcado que determina a linguagem B (“-”), visto que esta é um subconjunto próprio da linguagem A” (RAPOSO, 1992, p. 60). Então, se a criança nasce em uma comunidade em que se fala a língua B, a ligação primitiva do parâmetro será constituída por dados linguísticos primários, positivos, a que a criança terá acesso. Por outro lado, se essa criança tivesse nascido em uma comunidade onde se fala a língua A, ela teria acesso a dados positivos, que seriam de frases de A e não de B, considerados, então, inconsistentes com o valor inicial do parâmetro, mas que lhe

permitiriam com facilidade efetuar a mudança do valor de “-” para “+”, o que iria determinar a língua A.

Ao contrário desse princípio, supondo que o valor inicial não-marcado do parâmetro fosse o determinante da língua A, ou seja, a língua maior, e a criança crescesse numa comunidade falante dessa língua, a ligação inicial seria consistente com os dados primitivos, sendo, desse modo, mantida. Entretanto, se a criança cresce em uma comunidade onde se fala a língua B, haveria problemas. Isso aconteceria porque ela teria acesso a dados primários que são consistentes com a fixação inicial do parâmetro, já que a linguagem B é um subconjunto da linguagem A e, por isso mesmo, é parte integrante desta e determinará essa fixação. A criança necessitaria então da informação negativa de sua língua para que fosse efetuada a mudança do valor “+” para o valor “-”, ou seja “da informação de que as frases em A mas não em B não são gramaticais na língua (B) que está a adquirir” (RAPOSO, 1992, p. 60). Sendo assim, “O Princípio do Subconjunto, proibindo que o valor inicial do parâmetro seja aquele que determina a linguagem A, é de facto a formulação da hipótese de que a informação negativa não desempenha nenhum papel na aquisição da gramática” (RAPOSO, 1992, p. 61). O valor não marcado, ou seja, a posição *não-pro-drop*, assumida para o parâmetro do sujeito nulo, é, então, uma consequência direta do Princípio do Subconjunto. Isso significa que, em uma língua *pro-drop*, a omissão ou não do sujeito é opcional, enquanto em uma língua *não-pro-drop*, ela não é permitida, verificando-se sempre a realização do sujeito fonético. Assim, cada frase de uma língua *não-pro-drop* terá duas frases correspondentes em uma língua *pro-drop*. Para a seguinte frase, em inglês, por exemplo, que é uma língua *não-pro-drop*, “They arrived late”, há no português, língua *pro-drop*, duas possibilidades de realizações: “Eles chegaram tarde” e “Chegaram tarde” (RAPOSO, 1992, p. 61). Desse modo:

Qualquer língua *pro-drop* é assim “maior” que uma língua *não-pro-drop*, incluindo-a propriamente. Logo, o Princípio do Subconjunto determina que o valor inicial do parâmetro do sujeito nulo é a posição *não-pro-drop*, aquela que corresponde à linguagem ‘menor’ (RAPOSO, 1992, p. 61).

Se o Princípio do Subconjunto for atribuído à estrutura cognitiva da criança surge, desse modo, uma hipótese imensamente forte acerca do mecanismo de aquisição inicial e, particularmente, sobre o processo da aquisição, o que, de acordo

com Raposo (1992, p. 61), “é possível pôr à prova através de estudos empíricos sobre o desenvolvimento linguístico da criança”.

Caso as crianças tenham acesso a dados negativos de forma indireta, as circunstâncias em que isso ocorre deverão ser especificadas. De acordo com Lightfoot (1991), o acesso a dados negativos não implica menos riqueza na GU. Por outro lado, não são também muito fortes os argumentos para acesso a dados negativos, como afirma Lightfoot (1991, p.12):

So far, the arguments for indirect access to negative data are not very strong. Certainly, there are plausible alternative explanations for the null-subject phenomenon. One possibility is that the English setting for the relevant parameter is *unmarked*, i.e., the default case. Thus, Italian and Spanish children need specific evidence to adopt the other setting, and (12a) is the required evidence. The fact that the Italian setting for the parameter seems to be much more common across languages than the English setting does not entail that it is less marked, since markedness values do not reflect statistical frequency. In fact, Berwick's (1985) Subset Principle predicts that the Italian setting should be marked. The Subset Principle requires children to "pick the narrowest possible language consistent with evidence seen so far" (p. 237). The Italian setting of the parameter entails a language that is broader than one with the English setting (because in Italian subjects may or may not be phonetically expressed), and therefore the English setting must be unmarked (p. 290).<sup>33</sup>

Outra opção, de acordo com Lightfoot (1991), é que a variação em sujeitos nulos venha a se tornar dependente de alguma outra propriedade. O que se tem sugerido muito frequentemente é que a ocorrência de sujeitos nulos seja uma característica exclusiva de gramáticas com uma flexão verbal rica, o que em si só não é uma condição suficiente para que isso venha a acontecer. Conforme afirmação de Lightfoot (1991), embora uma língua como o alemão tenha flexão verbal que envolva número, pessoa e gênero, isso não é condição suficiente para que tenham sujeitos nulos, como ocorre com o espanhol, por exemplo. Desse modo,

---

<sup>33</sup> Até agora, os argumentos para acesso indireto a dados negativos não são muito fortes. Certamente há plausíveis explicações alternativas para o fenômeno do sujeito-nulo. Uma possibilidade é que a configuração inglesa para o relevante parâmetro é não marcada, isto é, o caso padrão. Assim, crianças italianas e espanholas precisam de evidências específicas para adotar outra configuração e (12a) é a evidência necessária. O fato de que a configuração italiana para o parâmetro pareça ser muito mais comum entre as línguas do que a configuração inglesa não implica que seja menos marcada, uma vez que os valores de marcação não refletem a frequência estatística. De fato, o Princípio de Subconjunto de Berwick (1985) prevê que a configuração italiana deve ser marcada. O Princípio do Subconjunto exige que as crianças "escolham o idioma mais estreito possível consistente com as evidências vistas até agora" (p. 237). O cenário italiano do parâmetro implica uma linguagem que é mais ampla do que com o cenário inglês (porque em italiano os sujeitos podem ou não ser foneticamente expressos), portanto, a configuração inglesa deve ser não marcada (p. 290) (tradução nossa).

quanto ao problema da aprendizagem, acrescenta ainda que esse permanece constante e que não é afetado pela riqueza das flexões.

Hyams (1983) fala sobre a impossibilidade de sujeitos nulos quando se trata de pronomes expletivos, como nas expressões *No, It's not raining / It's not cold out / There's no more / There's no money*<sup>34</sup> (Cf. Hyams, 1983, p. 136), que ocorrem em inglês, francês e alemão, mas não em italiano ou espanhol. Sobre essa questão, Hyams (1983) reuniu evidências que favorecem essas informações, ao se considerar os estágios de desenvolvimento por que passam as crianças, afirmando que gramáticas infantis, nos estágios iniciais, possivelmente teriam sujeitos nulos, sendo que o parâmetro tem o valor não marcado no italiano, o que violaria o Princípio do Subconjunto de Berwick (1982).

Para entendermos melhor a questão relacionada ao parâmetro do sujeito nulo no português do Brasil, abordaremos esta questão na próxima seção.

## 2.4 O SUJEITO NULO NO PORTUGUÊS DO BRASIL

A sintaxe do PB tem passado por profundas modificações em sua estrutura, o que a tem distanciado, cada vez mais, da sintaxe do PE. No que tange em relação às diferenças entre o PB e o PE, Galves (1987) aborda um dos aspectos, objeto de estudo do nosso trabalho, que é o sujeito nulo, mostrando também o uso do pronome *e/le* em posição sujeito, que, nesse caso, não tem o mesmo uso no PB e no PE: “Em PB é utilizado de preferência ao sujeito nulo, sem valor particular; em PE, fora os casos em que serve a desambiguar uma concordância idêntica em 1ª e 3ª pessoa, seu uso é contrastivo, correspondendo mais ao lui do que ao il francês” (GALVES, 1987, p.33).

Por outro lado, já nas relativas essa diferença vai assumir um caráter mais drástico, conforme Galves (1987), sendo que no PB, *e/le* na posição sujeito é geralmente utilizado como pronome “lembrete”, já no PE, esse emprego é agramatical, pois só é admitido na posição do sujeito das relativas uma “categoria vazia”, conforme exemplo abaixo, extraídos de Galves (1987, p. 33) e adaptados:

---

<sup>34</sup> Não, Não está chovendo / Não está frio lá fora / Não há mais / Não há mais dinheiro (tradução nossa).

Eu tinha uma empregada que ela respondia ao telefone e dizia... (NURC)  
(PB)

Eu tinha uma empregada que respondia... (PE)

Galves (1987) ainda acrescenta mais uma diferença, que é o uso constante de *e/le* em PB, posterior ao sintagma nominal sujeito, conforme exemplo abaixo de Pontes (1981 apud GALVES, 1987, p. 33):

(3) Essa competência ela é de natureza mental.

(ex. de E. Pontes)

Essas três diferenças, segundo Galves (1987) têm uma grande coerência entre si, sendo que no PE, *e/le* é marcado, enfático, não podendo servir, portanto, de retomada anafórica de um sintagma nominal. Já no PB, *e/le* é praticamente obrigatório para expressar a 3ª pessoa, por isso mesmo *e/le* pode retomar um SN. Por outro lado, “Deve-se notar, porém que essas duas características não são forçosamente ligadas nas línguas: il em francês não pode aparecer em posição sujeito de uma relativa apesar de ser obrigatório como marca de 3ª pessoa” Galves (1987, p. 33). Essa questão já havia sido enfatizada por Pontes (1981) como um funcionamento particular no PB, de acordo com que estaria relacionado com o fenômeno de língua de tópico, conforme veremos mais adiante.

Aos diferentes usos de *e/le* sujeito, irão corresponder também diferentes interpretações do sujeito nulo no PB e no PE, conforme afirma Galves (1987), já que os diferentes modos de interpretação dos sujeitos nulos têm relação ainda com os diferentes usos do *se*, pondo em jogo a alternância determinação/indeterminação. Sendo assim, reproduzimos o quadro abaixo, adaptado de Galves (1987, p. 37), que demonstra a alternância zero com *e/le* em PB e com *se* em PE:

	PE	PB
• determinado	usa saia	ela usa saia
• indeterminado	usa-se saia	usa saia

Desse modo, não se deve concluir que o sujeito nulo sempre receberá interpretação indeterminada em PB, de acordo com Galves (1987), sendo que a interpretação determinada ou indeterminada dependerá do contexto. Em virtude disso, Galves (1987) cita situações como *João disse que viria*, em que os falantes

brasileiros vão interpretar o sujeito de *viria* como sendo *João*. “Diremos que João controla referencialmente a categoria vazia sujeito do verbo subordinado” (GALVES, 1987, p. 37). O mesmo não ocorre para os falantes portugueses, já que essa interpretação não se torna obrigatória, pois o sujeito vazio em *João disse que viria*, “pode receber uma interpretação determinada independentemente da presença de um ‘controlador’” (GALVES, 1987, p. 37).

Outra situação diz respeito ao caso das infinitivas que indicaria novos contrastes na interpretação das categorias vazias, conforme exemplos de frases que, documentadas em PB, seriam agramaticais ou receberiam outra interpretação em PE, citadas a seguir, retiradas de Galves (1987, p. 37):

- (19) Eu estava pretendendo acabar antes de [e] nascer.  
 (20) Você quer deixar [e] ou [e] ficar com você?  
 (21) Você acha melhor [e] ir?<sup>35</sup>

Desse modo, então, o que ocorre na situação apresentada é:

Em (19) o sujeito de nascer é interpretado, no contexto, como o bebê do qual se falava; em (20), tanto o objeto de deixar quanto o sujeito de ficar é o texto do qual se fala. Em PE, (19) só poderia ser interpretado com o sujeito de nascer idêntico ao sujeito da sentença eu (a sentença seria assim gramatical, mas muito estranha) e (20) é definitivamente agramatical, o verbo querer impondo que o sujeito de ficar e o seu sejam idênticos (GALVES, 1987, p. 37).

Nesses casos, no PB é observado que o tópico da sentença vai competir com o sujeito para atribuir referência à categoria vazia, sendo que: “O mesmo se verifica em (21) em que o sujeito de ir pode ser você, mas pode também remeter a alguém de quem se fala” (GALVES, 1987, p. 37-38). Já no caso do PE, essa interpretação seria impossível. Desse modo, apresentamos o quadro, que reproduzimos abaixo, adaptado, indicando a interpretação do sujeito nulo nas infinitivas, conforme Galves (1987, p. 38):

PE - antecedente na sentença ou interpretação indeterminada  
 PB - antecedente no discurso ou antecedente na sentença (tópico)

<sup>35</sup> [e-] representa a categoria vazia, utilizado por Galves (1987).

Há, então, uma distinção nas construções infinitivas entre o PB e o PE, em se tratando de a possibilidade do sujeito nulo do infinitivo ser interpretado como se referindo à pessoa ou à coisa da qual se fala, de acordo com Galves (1987). No caso do PE, isso só é possível com a presença do pronome lexical, conforme exemplos reproduzidos abaixo de Galves (1987, p. 38):

- (22) Eu estava pretendendo acabar antes dele nascer.  
 (23) Você acha melhor ele ir.

Para que a frase seja aceitável em PE, nesse caso agora, não é a presença de um sujeito pronominal, conforme Galves (1987), mas a marca de pessoa, que tanto pode ser na 1ª e na 2ª, através da concordância, segundo exemplificado abaixo:

- (24) Eu estava pretendendo acabar antes de nascerem.  
 (25) Você acha melhor irmos? (GALVES, 1987, p. 38)

Sendo assim:

O que se depreende do contraste entre (19) e (20) por um lado, e (22) e (23) por outro lado, em PE, é que nesta língua, a 3ª pessoa não é interpretada como determinada no infinitivo. Se o pronome lexical não aparecer, o sujeito do verbo ou bem terá um antecedente na frase ou será interpretado como indeterminado (GALVES, 1987, p. 38).

O mais importante é observar que a mesma falta de determinação em relação à 3ª pessoa também será encontrada no PB nas frases com tempo finito, como afirma Galves (1987), já que a presença do pronome lexical *ele/ela* vai se tornar praticamente obrigatória como na situação citada previamente, quando há um antecedente possível como em *João disse que viria*. Por isso:

Nos dois casos, diremos que a flexão de 3ª pessoa é defetiva em relação ao traço /+pessoa/ que é o que assegura a interpretação determinada do sujeito nulo. E nos dois casos, a interpretação é de indeterminação ou de controle referencial por um antecedente (GALVES, 1987, p. 39).

O que será de primordial importância para diferenciar o PB do PE é a condição desse antecedente na sentença. No caso do PE, só há possibilidade de ser um sintagma nominal com função sintática na oração principal, de acordo com

Galves (1987), enquanto no PB, pode, além disso, ser também o tópico da sentença, isto é, o objeto ou a pessoa a respeito de quem a sentença é enunciada.

Como podemos observar, nesses estudos apontados por Galves (1987) são muitas as diferenças entre o PB e o PE, o que não pode se deixar de enfatizar, já que é algo bastante discutido, sempre que se fala em estudos que se referem à língua portuguesa. No entanto, vale salientar, que a questão do uso do sujeito pronominal, nessas duas variedades dialetais do português, se tornou um dos aspectos mais estudados e, principalmente, conforme já citado acima, o objeto de estudo de nossa pesquisa, que é o sujeito nulo.

Retomando de forma resumida o que foi citado acima por Galves (1987), vemos que, na posição sujeito, o pronome *e/le* não tem o mesmo uso no PB e no PE, sendo que no PB sua utilização se dá em preferência ao sujeito nulo. Nas relativas, o pronome *e/le* é utilizado como pronome lembrete, enquanto no PE, este uso é agramatical. O pronome *e/le* também é encontrado, no PB, posterior ao sintagma nominal, segundo Pontes, (1981), já no PE, *e/le* não pode servir de retomada anafórica de um sintagma nominal. Galves (1987) afirma ainda que os diferentes modos de interpretação dos sujeitos nulos têm relação com os diferentes usos do *se*, o que põe em jogo a alternância determinação/indeterminação. Cita também o caso das infinitivas em que frases documentadas em PB seriam agramaticais ou receberiam outra interpretação em PE.

Todas essas questões já foram abordadas e devidamente exemplificadas. Apenas fizemos esse apanhado do que foi apresentado aqui por Galves (1987), que é apenas um recorte, para enfatizar que essas e outras diferenças vêm nos mostrar o quanto PE e o PB se distanciam e como estudos como o que apresentamos aqui são relevantes na contribuição de buscas para se entender melhor a formação do português brasileiro e, mais precisamente, do português popular do Brasil.

Outro ponto importante é que a flexão verbal no PB está reduzida a quatro formas, como consequência da predominância do uso do “você”, em substituição ao pronome “tu”, de segunda pessoa do singular, conforme Galves (2001 [1987]), o que motivou também o uso do “vocês” na terceira pessoa do plural, o que voltaremos a abordar mais adiante. Temos, então, de acordo, com a tabela reproduzida de Galves (2001 [1987], p. 168):

	Singular	Plural
1ª pessoa	eu fal-o	nós fal-amos
2ª/3ª pessoas	você/ele fal-a	vocês/eles fal-am

Dessa forma, esse paradigma pode ser representado através da combinação de dois traços, positivo e negativo, de acordo com a seguinte configuração, reproduzida de Galves (2001 [1987], p. 168):

+pessoa/-número	-o
-pessoa/-número	-a
+pessoa/+número	-amos
-pessoa/+número	-am

Sendo assim, enquanto o traço do PB, por ser binário, possui semelhança de um traço sintático, por outro lado, o do PE, de acordo com Galves (2001 [1987]), tem como marca a oposição entre as três pessoas. Assim temos: 1ª pessoa, -o; 2ª pessoa, -as; 3ª pessoa, -a. Importante salientar a observação a respeito da “pobreza” e da “riqueza” da concordância, segundo Galves (2001 [1987], p. 168), que considera “o primeiro traço como definidor de uma concordância pobre, e o segundo, de uma concordância rica”.

Ainda sobre as diferenças entre o PB e o PE, encontramos um texto considerado pioneiro, em que Tarallo e Kato (2007 [1989]) proclamaram o que seria uma interface entre gerativismo e a sociolinguística, já citada na seção anterior e se refere ao que estamos buscando neste trabalho de pesquisa, que indica um caminho a ser percorrido, em relação a esse estudo. Desse modo, um panorama da sintaxe brasileira foi traçado por Tarallo (2018a [1993]), enfatizando as mudanças sintáticas ocorridas na virada do século XIX para o século XX. A afirmação de Tarallo (2018a [1993]) de que haveria uma assimetria entre os sujeitos expressos e o objeto nulo no português brasileiro (PB) motivou o interesse para as pesquisas sobre a realização dos sujeitos referenciais definidos. Através de um conjunto de fenômenos sintáticos, Tarallo (2018a [1993]) esboça a emergência de uma gramática brasileira, que mostrava uma diferença bastante clara em relação à gramática portuguesa,

apresentada através de quatro grandes mudanças sintáticas, a saber: (i) a reorganização do sistema pronominal e, conseqüentemente, a implementação do objeto nulo no sistema linguístico do PB, por um lado, e a realização mais frequente, por outro lado, do sujeito lexical; (ii) a mudança sintática havida nas estratégias de relativização como consequência direta da mudança no sistema pronominal; (iii) a reorganização dos padrões sentenciais básicos (ordem SVO); e (iv) a alteração nos padrões sentenciais em perguntas diretas e indiretas.

Essas diferenças entre o PE e o PB levaram a discussões sobre uma possível mudança do PB de uma língua de sujeito nulo, ou *pro-drop*, para uma língua não-sujeito-nulo, ou *não-pro-drop*. Apesar dos primeiros estudos, em relação às diferenças e semelhanças da sintaxe do sujeito nas variedades do PE e o PB tenderem a uma consideração sobre uma possível transição do português brasileiro para um estágio de língua *não-pro-drop*, ainda não há restrições que possam ser consideradas tão categóricas em relação ao licenciamento do sujeito nulo, como as observadas em línguas como o inglês, por exemplo. Enquanto Duarte (1995) defende a hipótese de que o português brasileiro segue lutando em direção ao sujeito pleno, ao citar como exemplo o francês medieval, que, como vimos, conviveu em torno de 150 anos com um sistema de sujeitos nulo até o parâmetro finalmente vir a se fixar como *pro-drop*. Por outro lado, Cyrino, Duarte e Kato (2000) chegaram à conclusão que a relação entre a referencialidade do sujeito e do objeto, nos processos de mudança envolvendo a pronominalização, era bastante significativa, conforme veremos mais adiante.

Por outro lado, há a proposta apresentada por Modesto (2004), ao afirmar que o PB não seria uma língua *pro-drop*, mas uma língua com parâmetro orientado para o discurso. Modesto (2004) não descarta a hipótese de o parâmetro do sujeito nulo estar diretamente relacionado à morfologia verbal “rica”, no entanto segue, para o PB, a mesma análise que Huang (1984) propôs para o chinês:

Vou tomar a posição de que, realmente, *pro-drop* está esfritamente relacionado à riqueza da flexão, seja qual for a definição de “rico”, e, então, línguas como o chinês, o finlandês e o PB não pertencem ao grupo de línguas de sujeito nulo, apesar das aparências. Vou argumentar, seguindo e em certo sentido fortalecendo a proposta de Huang (1984), que a presença de argumentos nulos nessas línguas está relacionada a um parâmetro distinto do parâmetro do sujeito nulo: um parâmetro que distingue línguas orientadas para a sentença e línguas orientadas para o discurso, que foi originalmente discutido por Tsao (1977) (MODESTO, 2004, p. 1-2).

Temos ainda a proposta de Barbosa, Duarte e Kato (2005) sobre o que se tem denominado de “competição de gramáticas”, o que provavelmente indicaria que há a possibilidade de que falantes do português brasileiro de futuras gerações se aproximem de um padrão mais semelhante ao do francês.

O que foi proposto por Cyrino, Duarte e Kato (2000), citado anteriormente, em relação a uma hierarquia referencial, é explicado, então, por Duarte (2012, p. 12):

Elementos com o traço [+humano] ocupariam o ponto mais alto da hierarquia, enquanto elementos não argumentais (como sujeitos expletivos) ocupariam o ponto mais baixo. Assim, os pronomes de 1ª e 2ª pessoas, inerentemente humanos, se encontram no ponto mais alto, enquanto a 3ª pessoa, que pode exibir o traço [±animado], [±específico], estaria num ponto mais abaixo; uma proposição (uma estrutura oracional ou uma porção maior do discurso) estaria ainda mais abaixo, enquanto um não argumento (como os expletivos) estaria no extremo oposto da hierarquia referencial.

A conclusão a que essas autoras chegaram é de que, na observação dos dois processos de mudança em curso do PB, ambos são guiados pela mesma hierarquia de referencialidade, mesmo que pareçam ter motivações independentes. Enquanto o objeto nulo foi implementado a partir da menor referencialidade dos antecedentes, o caminho inverso foi seguido pelos sujeitos pronominais plenos, que se implementaram a partir dos itens mais referenciais, segundo afirma Duarte (2012, p. 12-13):

O objeto nulo se implementou no sistema a partir dos antecedentes com menor referencialidade, exatamente os que têm como antecedente uma proposição, para só mais tarde atingir os antecedentes representados por um SN, iniciando pelos que exibiam o traço [-específico]/[-animado], e embora tenha mais tarde chegado a ocorrer com antecedentes com o traço [+específico]/[+animado], tem nesse tipo de referente seu ponto de resistência (Cf. Cyrino, 1993; 1997).

Os sujeitos pronominais plenos, seguindo caminho inverso, começaram a se implementar a partir dos itens mais referenciais – os pronomes de 1ª e 2ª pessoas, atingem mais lentamente a 3ª pessoa, a começar pelos antecedentes com o traço [+humano], e progredem mais lentamente com referentes de 3ª pessoa (Cf. Duarte, 1993).

Duarte (2012), ao analisar a trajetória do sujeito nulo no português do Brasil, utilizando-se de uma amostra de peças escritas e ambientadas no Rio de Janeiro, que representavam a fala de grupos sociais urbanos, em diferentes épocas, pode observar a ocorrência de uma mudança favorecendo o uso dos sujeitos pronominais expressos.

Outra questão que merece destaque, quando se fala no Parâmetro do Sujeito Nulo no PB são as considerações a respeito de línguas como o japonês, por exemplo. Conforme Kato (2018), o japonês e o Português Europeu (PE) são línguas com homofonia total, que licenciam sujeitos nulos referenciais e não-referenciais, mesmo sendo aparentemente bem diferentes no que diz respeito à identificação do sujeito nulo não referencial, já que o primeiro é identificado através do antecedente e o segundo através da concordância. Línguas como o japonês têm sido denominadas de línguas de sujeito nulo radical, já no caso de línguas como o português europeu de línguas de sujeito nulo consistente:

Na verdade, o japonês tem o mesmo mecanismo de identificação do sujeito nulo quando, no PE, temos a terceira pessoa.

(24) a. *O Pedro<sub>i</sub> está dizendo [que O<sub>i</sub> gosta de pizza].* PE

b. *Pita-wa [ O<sub>i</sub> pizza ga suki-da] to yute-iru.* Jap

Já as línguas sem homofonia, isto é [-SN] como o inglês, não admitem controle em ambiente finito, exigindo a repetição do sujeito referencial:

(25) *Peter is saying [ that he likes pizza]* (KATO, 2018, p. 150-151).

Ao manifestar o sujeito referencial, o PB vernacular tem um comportamento similar ao inglês, conforme afirma Kato (2018), demonstrando uma mudança em progresso para uma língua de sujeito nulo parcial, já que ocorre o sujeito nulo não-referencial em qualquer contexto: *Pedro<sub>i</sub> está dizendo [que ele<sub>i</sub> gosta de pizza]* (PB); *O Pedro disse [que O está chovendo]* (PB / PE) (Cf. Kato 2018, p. 151). Desse modo, o PB apresenta semelhanças de línguas como o islandês, denominadas *semi-pro-drop*, caracterizadas pela presença de sujeitos pronominais, quando referenciais, e nulos, quando não referenciais, segundo afirma Kato (2018). Sendo assim, o PB não respeitaria a característica do Princípio “Evite Pronome”, de Chomsky (1981), já que o pronome sujeito se manifesta em qualquer pessoa e não apenas em ambiente controlado (*Eu acho que eu passaria por causa da base que eu tinha. / Vocês dizem isso porque vocês são jovens. / Ela ficou solteira porque ela quis.*) (KATO, 2018, p. 15). Já em relação aos não referenciais, de acordo com Kato (2018), o PB não só tem expletivos nulos, como também tem sujeitos indefinidos e genéricos nulos, assim como o islandês, que é uma língua *semi-pro-drop*.

Em relação a essa questão, temos que levar em conta também afirmações anteriores de diversos autores de que o português é uma língua *pro-drop* parcial, como Holmberg, Nayudu e Sheehan (2009, p. 2):

Impressionistically speaking, null subjects in partial null-subject languages are optional in some contexts where they are obligatory in consistent null-subject languages and excluded in non-null-subject languages, and excluded in some contexts where they are allowed in consistent null-subject languages<sup>36</sup>.

Por outro lado, de acordo com Galves (2019), o sujeito nulo no PB tem sido bastante estudado, sendo que há uma concordância de todos os autores de que sua distribuição mais restrita do que no PE se deve ao empobrecimento da morfologia:

A questão de fundo, que divide as análises, é a de saber se a língua se caracteriza como uma língua *pro-drop*, ou seja, se a categoria vazia que ainda é legitimada em posição sujeito em certos contextos é a categoria pronominal *pro* (GALVES, 2019, p. 132).

Vários autores, no Programa Minimalista, de acordo com Galves (2019), defenderam a ideia de que o PB deixou de ser uma língua de sujeito nulo pronominal, isto é, uma língua *pro-drop*. Por outro lado, segundo Galves (2019), apesar do parâmetro do sujeito nulo ter um lugar de destaque em se tratando da teoria de parâmetros, o estudo ampliado de outras línguas mostrou que o a propriedade “sujeito nulo” estava longe de ser um fenômeno único, o que provocou um grande desafio ao programa de pesquisa associado à própria noção de parâmetro:

Vários autores analisam línguas que caracterizam como “sujeito nulo parcial”, das quais faz parte o PB. Essas línguas têm duas propriedades notáveis. A primeira é que o sujeito nulo é contextualmente mais restrito do que nas línguas canônicas de sujeito nulo, que passam a ser chamadas de línguas de “sujeito nulo consistente” (GALVES, 2019, p. 139).

De acordo com Kato (2021), a substituição da segunda pessoa “tu” pela forma de tratamento “Vossa Mercê”, considerado um paradigma clássico do Português, posteriormente reduzido a “você”, que, apesar de hoje ser considerado como pronome de segunda pessoa, exige concordância de terceira pessoa e mantém o caráter nominal de sua origem, como vimos no início dessa subseção. Essa mesma situação é observada em relação a segunda pessoa do plural “vocês”, que possui

---

<sup>36</sup> Falando de forma impressionante, os sujeitos nulos em línguas de sujeito nulo parcial são opcionais em alguns contextos em que são obrigatórios em línguas de sujeito nulo consistentes e excluídos em línguas de sujeito-não-nulo, e excluídos em alguns contextos em que são permitidos em línguas de sujeito nulo.

origem nominal em Vossas Mercês e, portanto, exige, do mesmo modo, concordância de 3ª pessoa plural. Também ocorre no sistema pronominal do PB outra mudança, de acordo com Kato (2021), que é a introdução do sintagma nominal “a gente”, expressão da 1ª pessoa plural, no entanto exige concordância de 3ª pessoa singular, já que esta é a sua natureza nominal.

Essas considerações a respeito da questão dos pronomes no PB vão levar ao que se tem propagado sobre o empobrecimento da morfologia de concordância, o que teria ocorrido devido à substituição de pronomes por nomes, modificando assim o sistema de referência, apesar da manutenção em algumas regiões do país, conforme Kato (2021, p. 258): “Tem-se, por exemplo, em algumas regiões do Brasil a manutenção do sistema pronominal, mas com a perda morfológica em algumas pessoas (“tu vai”; “nós vai”)”.

Seria, então, o empobrecimento da morfologia, resultante da modificação do sistema de referência pronominal, ocasionado pela substituição dos pronomes de 2ª pessoa, singular e plural, e 1ª pessoa plural por “você”, “vocês” e “a gente”, respectivamente, responsável no PB pela maior ocorrência do pronome sujeito em relação ao português europeu? O PB brasileiro estaria deixando de ser uma língua de sujeito nulo devido ao empobrecimento da morfologia de concordância?

Ao comentar sobre o sistema de referência das línguas, Kato (2021) afirma que eles são de dois tipos, de pronomes e de nomes, sendo que há línguas que podem ter os dois, ou seja, são línguas mistas. Dessa forma, de acordo com uma perspectiva paramétrica, conforme Kato (2021), línguas como o inglês e as línguas românicas, que dispõem de pronomes, são [+pronominais], enquanto as línguas asiáticas, como o japonês, são do tipo [-pronominal]. Por outro lado, temos então, o PB como uma língua de pronome parcial, do tipo [+/-pronominal]. Sendo assim, temos as seguintes classificações:

As línguas [+pronominais], [-pronominais] e [+/- pronominais] manifestam comportamentos distintos em sua morfossintaxe:

a) as línguas que usam pronomes em seu sistema de referência exigem concordância de pessoa e número no verbo:

Ex.: inglês (I am; you are, he is), espanhol (yo soy; tú eres, él es)

b) as línguas que escolhem nomes (formas de tratamento ou nomes próprios) têm concordância unipessoal e têm concordância apenas de 3ª pessoa, a que podemos designar de concordância 'default'.

Ex.: japonês (boku dessu 'sou eu'; kimi dessu 'é você'; kare-dessu 'é ele)

c) as línguas que podem dispor de um tipo misto com pronomes + nomes têm concordância empobrecida com a concordância dos impostores reduzida à 3ª pessoa.

Ex.: português brasileiro (PB)  
 Pronomes de 1ª p. singular (eu sou); de 3ª p. sing. (ele é);  
 Nomes: 2ª pessoa sing. (você é); 1ª pessoa plural. (a gente é) (KATO, 2021, p. 262).

Kato (2021) chama a atenção para a importância da distinção entre a pessoa semântica e a pessoa morfológica, bem como as consequências que essa questão gera no entorno sintático. Desse modo, é fundamental salientar que “você” e a “a gente” são, no contexto semântico, segunda pessoa do singular e primeira pessoa do plural do discurso, mas, por outro lado, quanto à concordância, ambos delineiam concordância de 3ª pessoa do singular (“você vai”; “a gente vai”) no seu entorno sintático.

A proposta apresentada por Kato (2021) é, então, a seguinte:

- a) línguas que dispõem de pronomes no seu sistema de referência são também línguas de Proeminência de Sujeito (exs.: inglês, espanhol, PE etc).
- b) línguas que dispõem de nomes, mas não de pronomes, são línguas de Proeminência de Tópico (ex.: japonês).
- c) Línguas que dispõem de pronomes e nomes em seu sistema de referência são línguas de proeminência de Tópico e de Sujeito (ex.: PB). (KATO, 2021, p. 264).

Além disso, Kato (2021) acrescenta, após as suas considerações, a concepção de sujeito nulo, de forma breve:

- a) Sujeito nulo em línguas de proeminência de sujeito é apagamento de pronome (HOLMBERG, 2005; ROBERTS, 2010);
- b) Sujeito nulo em línguas de proeminência de tópico é apagamento de nome (elipse) (TAKAHASHI, 2007, inter alia) (KATO, 2021, p. 267).

Para Kato (2021), o conceito de parâmetros enriqueceu e muito os estudos sobre a variação sintática, o que possibilitou, desde a sua implantação, um trabalho, cada vez mais intenso da teoria gerativa para mostrar as possíveis variações que as línguas apresentam. Desse modo, o Parâmetro do Sujeito Nulo tem sido um dos tópicos mais estudados.

Observando as hipóteses acima, temos que levar em conta que todas devem ser consideradas e que não podem ser descartadas, já que trazem análises consistentes e podem de certa forma, ser úteis em nossas conclusões sobre a formação do português popular brasileiro e, mais precisamente, sobre o que é

denominado de português afro-brasileiro, e nossos estudos sobre o sujeito nulo no PB. Não podemos descartar, no entanto, algumas hipóteses, que, de qualquer forma nos leva a um senso comum entre as que foram apresentadas pelos autores acima: (i) o PB não pode ser considerada uma língua *pro-drop* como o PE; (ii) o PB caminha para uma mudança para uma configuração como língua *não-pro-drop*.

As discussões acima vão nos levar para um ponto comum que é o fato do PB não se configurar como uma língua *pro-drop*. No entanto, a questão está em relação a qual hipótese deverá ser aceita como sendo considerada mais próxima do nosso trabalho de pesquisa. Vamos retomar essa discussão, voltando, de forma resumida ao que foi discutido acima. Como vimos, as diferenças entre o PE e o PB levantaram possibilidades sobre uma possível transição do português brasileiro para um estágio de língua *não-pro-drop*, apesar de haver restrições de que isso possa ser de forma tão categórica como no inglês, por exemplo.

Seguindo o que Huang (1984) propôs para o chinês, neste caso o PB seria uma língua de tópico discursivo, assim como acontece também com o finlandês. Em relação a esta questão, também temos Pontes (1981) e Galves (1987) afirmando que o PB se aproximaria mais das línguas tópico-comentários como o chinês, se distanciando de línguas como o inglês de sintaxe sujeito-predicado. Por outro lado, Modesto (2004), afirma que o PB não seria uma língua *pro-drop*, mas uma língua com parâmetro orientado para o discurso. Vimos também o que é afirmado por Barbosa, Duarte e Kato (2005) sobre o que se denomina “competição de gramática”, indicando a possibilidade das futuras gerações se aproximarem de um padrão mais similar ao francês. Já a hipótese de Kato (2018) é que o PB vernacular tem um comportamento semelhante ao inglês. Antes, porém, levantou considerações sobre o japonês e o PE, afirmando que línguas como o japonês são denominadas de sujeito nulo radical e o PE de sujeito nulo consistente. Kato (2021) também afirma que o PB pode ser considerada uma língua de pronome parcial, do tipo [+/- pronominal]. e possui apagamento de pronome e elipse nominal. Há ainda a afirmação anterior de diversos autores como Holmberg, Nayudu & Sheehan (2009) de que o português é uma língua *pro-drop* parcial.

A seguir, veremos algumas considerações a respeito do caráter da concordância variável sujeito-verbo e o parâmetro do sujeito nulo no PB.

#### 2.4.1 O caráter variável da concordância sujeito-verbo e o parâmetro do sujeito nulo no PB

Importante salientar que para se buscar respostas para as questões referentes a um processo de variação/mudança linguística, se faz necessária uma teoria. Desse modo, é possível associar o modelo da Teoria da Variação e Mudança Linguística, proposto por Weinreich, Labov e Herzog (2006 [1968]), com a Teoria de Princípios e Parâmetros, de Chomsky (1988 [1981]), principalmente quando se procura investigar os problemas de um processo de mudança em curso, como as restrições estruturais, a implementação, o encaixamento, a transição e a avaliação.

Desse modo, dentro da proposta do nosso trabalho de verificarmos a ocorrência do sujeito nulo no português afro-brasileiro na comunidade quilombola de Lagoinha, partimos desse pressuposto. A associação entre as teorias linguísticas contribui para a formulação de hipóteses que vão indicar quais os efeitos colaterais ou mesmo os vestígios resultantes da entrada de uma nova forma no sistema, o que para Weinreich, Labov e Herzog (2006 [1968]) seria ingênuo afirmar que não ocorra durante esse processo. Teorias linguísticas são necessárias para que se desenvolvam análises e interpretações de mudanças em curso.

Em se tratando da mudança que caminha para a utilização do sujeito pronominal de referência definida expresso, o mais provável é esperar que os sujeitos de menor referencialidade sejam também afetados, mesmo que de forma mais lenta. Diante disso, a Teoria de Princípios e Parâmetros fornece os meios para interpretar e estabelecer a relação entre estruturas que, aparentemente, são independentes. Para Roberts (1993), o PB contemporâneo se apresenta como um caso de mudança paramétrica em progresso, por isso mesmo, se faz necessário ampliar os estudos, o que seria muito relevante para a descoberta de propriedades fundamentais de um sistema defectivo de sujeitos nulos.

Em relação a esse processo paramétrico de mudança, Duarte (1993), buscando evidências que confirmassem a análise de Tarallo (1983), apresentou em seu trabalho, que teve como base uma amostra de textos de teatro escritos no Rio de Janeiro nos séculos XIX e XX, conforme já referido, resultados que indicam um processo de mudança em progresso no PB em direção ao uso preferencial por sujeitos pronominais expressos, sugerindo uma remarcação no valor do parâmetro do sujeito nulo. Essas mudanças em curso no PB podem e devem ser estudadas em

uma língua viva constituída por um patrimônio linguístico privilegiado, o que estamos buscando ao realizarmos o nosso trabalho de pesquisa linguística sobre o sujeito nulo no português afro-brasileiro. Nesse sentido:

Assumir uma análise empírica com pressupostos da teoria gerativa poderia parecer uma heresia, e, de fato, as reações iniciais ao artigo-manifesto de Tarallo (1987) e ao casamento teórico confirmado em Tarallo e Kato (2007 [1989]) foram veementes, mas tudo isso já foi há muito tempo exorcizado. A essa altura, ninguém pode imaginar que um modelo de estudo da mudança possa prescindir de uma teoria linguística, seja ela formalista ou funcionalista (DUARTE 2012, p. 14).

Desse modo, “O PB estaria passando de língua [+sujeito nulo] para língua [-sujeito nulo]” (DUARTE, 2012, p. 21). Uma das motivações para esse processo de mudança seria o fato de *você* ter sido inserido no quadro pronominal e a neutralização entre *você* e *tu*, além da presença do pronome *a gente* competindo com o pronome *nós*, o que contribuiria para a redução do paradigma verbal no PB:

Essa redução do paradigma flexional seria, pois, o resultado de entrada de novos pronomes, que se combinam com a forma verbal sem desinência distintiva e com o apagamento de algumas desinências, como a de segunda pessoa <-s> e a de terceira do plural <-m>, em certas variedades ou contextos. Como apontava Galves (1993), o PB tinha perdido a distinção do traço [+pessoa], restando-nos apenas a distinção de número (Duarte, 2012, p. 22).

Os resultados apresentados indicaram que havia um percurso diferente da mudança segundo a pessoa do discurso. Durante o período compreendido entre todo o século XIX e as primeiras décadas do século XX, foi observado que o PB, nos textos estudados, revela um comportamento padrão de línguas românicas de sujeito nulo, em que o sujeito expresso é a forma “marcada” e o sujeito nulo é o *default*:

Ao longo desses quase cem anos, o paradigma pronominal exibia *você* e *tu*, em distribuição complementar, com a primeira forma marcando um distanciamento em oposição à intimidade que caracterizava a segunda. Naturalmente, a forma de tratamento, o(a) senhor(a), é igualmente usada como forma de cortesia e nas relações assimétricas. Tal paradigma pronominal, que incluía formas verbais idênticas - ou sincretismos - para referência à segunda pessoa (chamada tradicionalmente de “indireta”, porque provinha de uma forma nominal, hoje plenamente gramaticalizada) e à terceira do singular, não constituía obstáculo ao licenciamento e à identificação de um sujeito nulo (DUARTE, 2012, p. 23).

Já no período que se segue, nas peças escritas a partir de 1930, Duarte (2012) comenta que ocorre a substituição do pronome *tu* por *você*, o que vai aumentar de forma expressiva o índice de sujeitos pleno de 2ª pessoa. Passa, então, o pronome expresso a ser a forma não marcada para a 2ª pessoa. No entanto, ela enfatiza que o pronome *tu* não desaparece por completo, sendo encontrado em algumas peças, ocorrendo uma alternância entre o *tu* e o *você* na fala das personagens. O *tu* vai se manter ausente nos textos analisados entre 1950 e 1970, voltando à escrita na última peça analisada, já na década de 1990. Nesse caso, o *tu* vai aparecer em variação com o *você*, apresentando a forma verbal zero na desinência e tendo o pronome expresso absolutamente.

Quanto à 1ª pessoa, o preenchimento do sujeito é consistente:

Em relação à primeira pessoa, a curva não é tão brusca, mas é consistente na direção do preenchimento do sujeito. Os índices de 31%, 21% e 17% são perfeitamente compatíveis com sistemas de sujeito nulo, em que um pronome de primeira pessoa é expresso para o falante se reintroduzir no discurso, como mostra Marins (2009) para o italiano oral, ou ainda para expressar ênfase ou contraste. É digno de nota o fato de que, entre os sujeitos de primeira pessoa expressos no primeiro período, quase metade aparece em construções do tipo V2, perdidas no PB (DUARTE, 2012, p. 24).

Nos anos 1990, quando surge *a gente* como forma pronominal nos textos teatrais, a mudança, acionada pela 2ª pessoa, tem continuidade na 1ª pessoa:

Apesar de ser a única forma verbal com desinência distintiva no singular, a segunda pessoa ganha no plural a concorrência entre o pronome *nós* (combinado com forma verbal de desinência distintiva <-mos>) e o pronome *a gente*, que, graças à sua origem nominal, se combina com a forma verbal não marcada ou de desinência zero (cf. Lopes, 2003) (DUARTE, 2012, p. 25).

Já a 3ª pessoa demonstra um comportamento diferenciado, exibindo índices próximos da 1ª pessoa nos quatro primeiros períodos (17%, 35%, 34% e 28%) e um aumento mais estável (41%, 48% e 45%), mesmo que paulatinamente, em relação às 1ª e 2ª pessoas, conforme Duarte (2012, p. 25): “Uma evidência de mudança qualitativa, porém, está no fato de que um sujeito nulo de terceira pessoa nos três últimos períodos é mais restrito estruturalmente do que nos períodos anteriores” Nas primeiras sincronias, a identificação do sujeito nulo é feita por um antecedente, que está em outra função sintática. Já a partir da segunda metade do século XX, a condição essencial para que isso ocorra, é a presença de um antecedente com a

mesma função, que tanto pode ser em uma oração que se encontra próxima, quanto no mesmo período.

A conclusão a que se chega é: “A redução no quadro de desinências verbais alterou as características de língua de sujeito nulo do PB” (DUARTE, 2012, p. 26). Sendo assim um sistema flexional “rico” seria responsável pelo licenciamento e identificação de um sujeito nulo, o que é comprovado pelo comportamento diferenciado da 3ª pessoa. Seria, então, necessário um reforço externo ao elemento da concordância e o caráter anafórico da 3ª pessoa retardaria o processo. Duarte (2012) aprimora a investigação do comportamento da 3ª pessoa, ampliando a amostra usada em Duarte (1993):

A presente análise focalizará dois aspectos em especial: o primeiro, de natureza estrutural, refinando o tratamento da “acessibilidade” do antecedente do sujeito, aspecto brevemente comentado na seção precedente e aqui referido como “padrão sentencial”, conforme proposto em Barbosa, Duarte e Kato (2005); o segundo, de natureza semântica, leva em conta a hierarquia referencial proposta em Cyrino, Duarte e Kato (2000), que chama a atenção para o papel da interação dos traços [+/-humano] e [+/-específico] nos processos de mudança em direção a pronomes expressos, que é o caso da implementação dos sujeitos pronominais, e nos casos de mudança em direção a uma categoria vazia, como ocorre com a implementação do objeto nulo (cf. Apresentação deste volume). Finalmente, a observação dos dados, levou à inclusão de um grupo de fatores ligado ao tipo de verbo, sob a hipótese de que, juntamente com o traço semântico [-hum/-esp] e a função diferente do antecedente, os verbos de ligação constituiriam um importante contexto de resistência do sujeito nulo no PB (DUARTE, 2012, p. 26-27).

De acordo com a proposta de Barbosa, Duarte e Kato (2005), foram analisados os padrões sentenciais. Neste caso, levou-se em conta a função do antecedente e a ocorrência deste no mesmo período ou em períodos próximos:

- i. Padrão 1: O antecedente se encontra no mesmo período e é o sujeito da oração precedente (principal ou subordinada);
- ii. Padrão 2: O antecedente se encontra no período adjacente e tem a função de sujeito ou tópico estrutural/discursivo;
- iii. Padrão 3: O antecedente é o sujeito de uma oração não adjacente no contexto precedente, ou seja, há uma ou mais orações intervenientes;
- iv. Padrão 4: O antecedente está na oração precedente/adjacente, mas tem função distinta da de sujeito (DUARTE, 2012, p. 27-29).

De acordo com a proposta de Barbosa, Duarte e Kato (2005), foram analisados os padrões sentenciais. Neste caso, levou-se em conta a função do antecedente e a ocorrência deste no mesmo período ou em períodos próximos:

- v. Padrão 1: O antecedente se encontra no mesmo período e é o sujeito da oração precedente (principal ou subordinada);
- vi. Padrão 2: O antecedente se encontra no período adjacente e tem a função de sujeito ou tópico estrutural/discursivo;
- vii. Padrão 3: O antecedente é o sujeito de uma oração não adjacente no contexto precedente, ou seja, há uma ou mais orações intervenientes;
- viii. Padrão 4: O antecedente está na oração precedente/adjacente, mas tem função distinta da de sujeito (DUARTE, 2012, p. 27-29).

Em sua análise, Duarte (2012) levantou 1.145 dados, sendo que a variável dependente foi o sujeito pleno, incluindo diferentes sincronias em um grupo de fatores. O padrão sentencial, o traço semântico/referencial do referente, o período de tempo e a transitividade verbal foram selecionados como relevantes para a realização da variável. O que favoreceu o sujeito expresso foi a distância do antecedente e a função distinta da de sujeito. Sendo assim, esses contextos indicam por onde a mudança se implementará com maior facilidade. Em relação ao traço [+h] para o preenchimento do sujeito, combinado com o traço [+esp], foi encontrado o peso 61, indicando a relevância da hierarquia referencial no preenchimento do sujeito.

Ao analisar o tempo em que os textos foram produzidos, o sujeito nulo de 3ª pessoa é mais favorecido no século XIX. Já na segunda metade do século XX, diante a uma mudança gradual favorecendo o sujeito pleno, com 53 anos (1955), 67 anos (1975) e 77 anos (1990), com uma diferença significativa superior a .10 entre dois pesos. Por isso mesmo:

Não há dúvida de que o padrão sentencial é um fator importante no processo de mudança gramatical. E tal mudança só pode ter sido desencadeada pela alteração no paradigma flexional. O fato de os padrões 1 e 2 serem os mais resistentes se deve, naturalmente, à acessibilidade do referente (DUARTE, 2012, p. 35).

Já em relação à análise do traço semântico do referente foi encontrado um número reduzido de dados para o traço [-específico], confirmando o papel da hierarquia referencial na implementação do sujeito expresso. Nesse caso, destaca-se o traço [+humano] juntamente com o traço [+específico]. Desse modo, conforme Duarte (2012, p. 37), “O traço [-humano] se revela um inibidor do pronome expresso, seja combinado com o traço [+esp], com ocorrência total bastante expressiva (152 dados), seja com o traço [-esp], com 48 dados”. Por outro lado, a expressão do

pronome sujeito é totalmente desfavorável em um contexto em que aparece a combinação dos traços [-hum] com o traço [-esp].

Sendo assim:

Aliando os resultados para o traço semântico à análise dos padrões sentenciais, é possível concluir que os contextos que favorecem a implementação do pronome pleno são de natureza estrutural e semântica. Confirma-se, assim, a relevância do padrão sentencial, em que fatores de natureza funcional, como a maior ou menor acessibilidade do antecedente, estão em jogo, com um fator referencial, que independe do padrão sentencial e confirma a atuação da hierarquia proposta em Cyrino, Duarte e Kato (2000), que guia um processo de mudança envolvendo a expressão de pronomes (DUARTE, 2012, p. 39).

Duarte (2012) comparou a pesquisa de Duarte (1995) com os resultados obtidos com a análise das peças de teatro, mostrando resultados muito parecidos, principalmente no que se refere aos textos escritos nos anos de 1990 (Período VII). Em ambos os resultados, o preenchimento do sujeito é maior na 2ª pessoa, seguido pela 1ª. Já na 3ª pessoa, o favorecimento do sujeito expresso se dá através do traço [+hum/+esp] (68% e 64%, respectivamente). As outras combinações [+hum/-esp], [-hum/+esp] e [-hum/-esp] revelam resistência ao preenchimento do sujeito, demonstrando índices inferiores. Desse modo, a conclusão de Duarte (2012) é de que o preenchimento do sujeito é bastante influenciado pelo traço [+humano] do antecedente, sendo que a associação do traço [-humano] com o traço [+específico] do antecedente tem diminuído a resistência ao longo do tempo. Sendo assim, o traço que mais favorece a resistência da implementação do sujeito expresso seria o [-específico].

Quanto à questão do enfraquecimento da concordância, que é outro aspecto da gramática do PB, de acordo com Galves (2019), se refere ao caráter variável da concordância sujeito-verbo, conforme exemplo abaixo:

- a) Os meninos(s) **jantou/jantaram**.
- b) As criança(s) **brincava/brincavam** na varanda Galves (2019, p. 143).

Segundo Galves (2019), os estudos em relação a esses assuntos são geralmente limitados ao contexto sociolinguístico ou histórico, enquanto esse tipo de variabilidade é considerado, em sua maioria, como resultado de um processo de criouliização pelo qual passou o português no Brasil. Por isso mesmo: “É muito raro ser encontrado em trabalhos de inspiração gerativista, pela dificuldade desse

modelo em integrar a variação” (GALVES, 2019, p. 143). Uma observação importante, em primeiro lugar, que deve ser tirada dos estudos quantitativos, de acordo com Galves (2019), está relacionada ao fato de a variabilidade da concordância perpassar todos os dialetos do PB, mesmo que seja de maneira bastante diferente:

Lucchesi, Baxter e Silva, com base em vários estudos sociolinguísticos, mostram que, na terceira pessoa do plural, a concordância entre o sujeito e o verbo se dá em 16% das ocorrências no dialeto afro-brasileiro de Helvécia e em 94% no dialeto de universitários cariocas. Entre esses dois polos, ainda no Rio de Janeiro, encontra-se 73% de concordância na fala de moradores escolarizados e enfim 38% na fala de pescadores do estado do Rio de Janeiro (GALVES, 2019, p.143).

Sendo assim, é possível associar a variabilidade da concordância, de acordo com Galves (2019), submetida evidentemente à pressão da norma urbana e escolar, ao enfraquecimento do paradigma verbal. Por outro lado, a forma usada nas ocorrências de não concordância é a de terceira pessoa do singular:

Na “primeira viagem”, argumentei que essa forma correspondia à combinação de traços [-pessoa]/[-plural], derivando disso a sua interpretação indeterminada. A terceira pessoa do singular no PB é assim totalmente desprovida de capacidade referencial. Isso sugere que na ausência de concordância verbal, quando o sujeito contém um traço de número, associado ou não a um traço de pessoa, esse sujeito funciona como ligador de flexão verbal (GALVES, 2019, p.143).

A concordância se encontra de fato no componente morfológico, sendo que é possível elaborar uma análise, segundo Galves (2019), articulando as propriedades da derivação sintática com as propriedades do paradigma morfológico sem transferir as segundas para as primeiras. Desse modo, a dificuldade de definição do limiar de riqueza/pobreza morfológica dos traços seria resolvida, segundo Galves (2019), como ficou comprovado nas dificuldades encontradas na formulação do parâmetro *pro-drop*. Por isso mesmo, o componente morfológico e a derivação sintática devem conspirar para uma efetiva realização das estruturas que estão associadas aos enunciados possíveis na língua:

Esse modelo é inteiramente coerente com um modelo no qual a sintaxe é dirigida pela morfologia na aquisição, e particularmente interessante para derivar a variação morfológica decorrente da situação de contato, da qual o PB nos dá interessantes efeitos (GALVES, 2019, p.145).

No entanto, se considerarmos o ponto de vista da gramática independentemente do processo de aquisição, a morfologia não dirige a sintaxe, já que a sintaxe não codifica a morfologia (GALVES, 2019, p. 145).

É importante salientar que o fato de estar ocorrendo um processo de mudança em um parâmetro da gramática implica uma mistura de duas marcações opostas. A mudança, então, ocorre lentamente ao longo do tempo. À medida que a mudança avança, é natural que exista essa competição entre as duas marcações durante esse processo. Por isso mesmo, a mudança de valor no português brasileiro, de positivo para negativo, demonstra essa competição, que é comum e ocorrerá ao longo do tempo, durante o avanço dessa mudança.

Como vimos, são muitas as propostas no que se refere aos estudos sobre o sujeito nulo no PB. Todas merecem a devida atenção, mas o nosso intuito ao apresentá-las é mostrar somente as muitas alternativas que advém dos estudos acerca do parâmetro do sujeito nulo no PB. No entanto, consideraremos a proposta de Duarte (1993), de que o PB não é uma língua *pro-drop* e, dessa forma, analisaremos a força das hipóteses a respeito das mudanças nos paradigmas pronominal e flexional como ocasionadoras da prevalência da realização lexical dos pronomes sujeitos no PB, a partir dos dados de fala dos informantes da comunidade quilombola de Lagoinha no município de Nova Canaã, na região Sudoeste da Bahia, conforme veremos mais adiante.

Outra questão que merece ser enfatizada e um dos fatores que teriam contribuído para a mudança de parâmetros no português brasileiro, de *pro-drop* para *não-pro-drop*, teria sido o contato entre língua durante a formação do português brasileiro ocorrida no período da colonização do Brasil, que também levaremos em consideração e é de fundamental importância na elaboração desta pesquisa. Segundo Lucchesi (2009b, 2015), o português adquirido de forma precária pelos africanos e pelos indígenas que se integraram à sociedade colonial, através de um processo de transmissão linguística irregular é o fator mais importante quando se fala nas atuais variedades da realidade linguística brasileira, o que será abordado logo a seguir.

A partir da próxima seção, abordaremos sobre a importância dos africanos na formação da sociedade brasileira, o plurilinguismo das línguas africanas no Brasil, a influência das línguas africanas na formação do PB e o que poderia ser configurado como Português Afro-brasileiro.

## 2.5 CONCLUINDO A SEÇÃO

Apresentamos nesta seção a Teoria de Princípios e Parâmetros, proposta por Chomsky (1981), fazendo também um breve relato dos estudos anteriores sobre a Gramática Gerativa, em que mostramos o percurso até a elaboração dessa teoria. A partir daí, abordamos o quanto os estudos gerativistas ganharam uma nova dimensão com a Gramática Universal composta essencialmente, conforme Chomsky (1988 [1981]), como um conjunto de regularidades gramaticais universais, denominados de princípios e, por outro lado, de um conjunto limitado de variações linguísticas possíveis, que são os parâmetros. Essa mudança do conceito anterior da GU de um sistema de regras para um modelo fundamentado em princípios e parâmetros é, de acordo com Chomsky (1994 [1986]), considerado de fundamental importância para os estudos linguísticos.

A partir dessa visão geral sobre a GU, passamos a abordar a questão dos parâmetros e dos gatilhos na aquisição da linguagem. Outro ponto também abordado é a faculdade da linguagem, enquanto componente da mente/cérebro, de acordo com Chomsky (1978 [1965], 1988 [1981]). Dessa forma, a GU é o estado inicial da faculdade da linguagem ( $S_0$ ) e a gramática do indivíduo adulto, o seu estado final ( $S_s$ ). Para que a aquisição da linguagem seja possível para o ser humano, os princípios e parâmetros são constituídos pela GU como uma faculdade inata e autônoma, que funciona independente de outros módulos da mente/cérebro.

Após essa discussão, discorreremos sobre um dos parâmetros mais discutidos na teoria gerativa, quando se fala da Teoria de Princípios e Parâmetros, que é o Parâmetro do Sujeito Nulo, objeto de estudo do nosso trabalho. Vimos, então, que as línguas humanas são compostas por sujeitos sintáticos e seus respectivos predicados, de acordo com um dos Princípios da GU, sendo que este é um dos princípios rígidos da GU: a existência, nas línguas humanas da posição de sujeito nas orações, o que não indica que essa posição tenha que ser preenchida obrigatoriamente por um NP com conteúdo fonético. Dessa forma, em se tratando de parâmetro, o Sujeito Nulo irá variar de forma binária entre as línguas, sendo que algumas línguas o marcarão como positivo e outras como negativo. Quando esse parâmetro é marcado como positivo numa língua, dizemos que ela é [+ sujeito nulo]. Quando é marcado como negativo, a língua é considerada [- sujeito nulo]. Chomsky

(1988 [1981]), então, afirma, sobre esta questão, que as propriedades relacionadas ao parâmetro *pro-drop* são o fator mais relevante quando se fala na teoria de princípios e parâmetros.

Finalizando a seção, abordamos o Parâmetro do Sujeito Nulo no PB, a partir da questão de que sintaxe do PB tem passado por profundas modificações em sua estrutura, o que o tem distanciado, cada vez mais, do padrão linguístico do PE (GALVES, 1987, 2019). Dando seguimento, vimos a afirmação de Tarallo (1983) sobre a assimetria entre os sujeitos expressos e o objeto nulo no PB, que resultou na motivação para a realização de pesquisas sobre a realização dos sujeitos referenciais definidos. A seguir, vimos os estudos de Cyrino, Duarte e Kato (2000) que nos levaram à conclusão de que a referencialidade do sujeito e do objeto era bastante significativa nos processos de mudança envolvendo a pronominalização no PB. Roberts (1993), então, comenta sobre esta questão e afirma que o português brasileiro contemporâneo se apresenta como um caso de mudança paramétrica em progresso, por isso mesmo, se faz necessário ampliar os estudos, o que seria muito relevante para a descoberta de propriedades fundamentais de um sistema defectivo de sujeitos nulos. Essas mudanças em curso no PB podem e devem ser estudadas em uma língua viva constituída por um patrimônio linguístico privilegiado. Já para Duarte (2012), o PB estaria passando de língua [+sujeito nulo] para língua [-sujeito nulo].

Voltaremos a falar e nos aprofundarmos mais sobre o Parâmetro do Sujeito Nulo no Português do Brasil, um pouco adiante, quando analisarmos os dados referentes à pesquisa deste trabalho sobre o português afro-brasileiro na comunidade quilombola de Lagoinha, na região sudoeste da Bahia. A seguir abordaremos sobre o português afro-brasileiro, a presença das línguas africanas em terras brasileiras e o português falado pelos africanos no período colonial brasileiro.

### 3 O PORTUGUÊS AFRO-BRASILEIRO

Ao falarmos das condições sociais que desencadearam as mudanças na estrutura linguística do português, tanto no Brasil quanto em Portugal, é de fundamental importância explorar o campo sócio-histórico que contribuiu para que tais alterações viessem a ocorrer já que, conforme Lucchesi (2009b), o contato dos colonizadores portugueses com um contingente de milhões de aloglotas, que já se encontravam em terras brasileiras, falantes de mais de mil línguas indígenas nativas, além de, aproximadamente duzentas línguas trazidas juntamente com cerca de quatro milhões de africanos escravizados é, imprescindivelmente, “o principal parâmetro histórico para a contextualização das mudanças linguísticas que afetaram o português brasileiro” (LUCCHESI, 2009b, p. 41).

Sendo assim, é necessário fazermos uma breve retrospectiva sobre a história sociolinguística do Brasil, iniciada com a vinda dos primeiros colonizadores portugueses que, ao chegarem, entraram em contato com os indígenas que habitavam a costa, no início do século XVI. Esses povos indígenas, de tribos que falavam línguas diferentes, se comunicavam entre si através de uma espécie de *koiné*. Os portugueses passaram a utilizar essa língua franca, denominada de *língua geral*, que foi usada para que os colonizadores se comunicassem com os índios do litoral e pudessem utilizá-los como mão de obra para a extração do pau-brasil e, posteriormente, para o cultivo da cana-de-açúcar, do tabaco e do algodão. Também os jesuítas utilizaram essa língua geral, adotando-a no seu projeto de conversão dos indígenas, incluindo aí a sua codificação nos moldes da gramática portuguesa da época. Sobre a língua geral, Lucchesi (2009b) apresenta uma diversidade de situações linguísticas referentes a esse termo:

- (i) a *koiné* empregada na comunicação entre as tribos de línguas do tronco tupi da costa brasileira;
- (ii) a sua versão como língua franca usada no intercurso dos colonizadores portugueses e indígenas;
- (iii) a versão nativizada predominante nos núcleos populacionais mestiços que se estabeleceram no período inicial da colonização; e
- (iv) a versão “gramaticalizada” pelos jesuítas sob modelo do português e utilizada largamente na catequese, até de tribos de língua não tupi – chamados por estes de tapuias, que significa ‘bárbaro’, em tupi (LUCCHESI, 2009b, p. 43).

Deve-se levar em conta também o povoamento do Brasil e a miscigenação que originou os mamelucos, filhos dos colonizadores portugueses com as índias que habitavam o litoral paulista. Os mamelucos eram exímios caçadores de índios, contribuindo para a exploração dos indígenas pelos bandeirantes, que através de expedições denominadas bandeiras adentraram os sertões de Minas Gerais, o que contribuiu ainda mais para a expansão da *língua geral*. A predominância da língua geral em São Paulo durou até o início do século XVIII. Entre os anos de 1560 e 1580, segundo Lucchesi (2009b), os tupinambás, fugindo dos colonizadores de diversas regiões do litoral brasileiro, deslocaram-se para o Maranhão e depois para as regiões banhadas por afluentes do Rio Amazonas.

Desse modo, as bandeiras foram responsáveis pela expansão da língua geral até entre tribos de língua não tupi. Encontra-se, então, mais uma situação em que se inclui o termo língua geral: “(v) língua franca de base tupi utilizada como língua segunda por tribos de língua não tupi (podendo nesses casos ocorrer a sua nativização)” (LUCCHESI, 2009b, p. 44). Assim sendo, a língua geral chegou também à Amazônia e foi adotada por tribos não tupi, transformando-se no *nheengatu*, que quer dizer “língua boa”, mantendo-se em uso até meados do século XX.

Com base no exposto anteriormente, a chegada do grande contingente de negros escravizados, que teria tido o seu apogeu no século XVII, para substituir a mão de obra indígena, é outro ponto relevante para o contato linguístico no Brasil. Devido à resistência ao trabalho forçado, além de a campanha dos jesuítas ser contra a escravidão indígena, essa teve que ser substituída para garantir o projeto da colônia agroexportadora de açúcar. Conforme Lucchesi (2009b), o tráfico negreiro foi autorizado oficialmente no Brasil por um alvará de D. João III, de 29 de março de 1549. Apesar disso, estima-se que a vinda dos primeiros negros escravizados já tenha ocorrido a partir da primeira expedição oficial com o intuito de povoamento do Brasil em 1532.

Nesse contexto, com a importação dessa população de negros vindos da África, Pernambuco e Bahia são os lugares do Brasil que recebem o maior número de escravos, prática que cresceu com muita rapidez, sendo que no final do século XVI a população negra já é maioria na base da sociedade colonial brasileira, aumentando ainda mais no século XVII. Por outro lado, a língua geral continuava predominando nas regiões mais periféricas da colônia (São Paulo e Maranhão) e no

interior do país, devido à aculturação das populações indígenas. Desse modo, ocorria o avanço da língua portuguesa através dos centros mais ligados à atividade mercantilista, como Bahia e Pernambuco, que integravam o projeto colonial do Brasil:

Nesses dois centros, e depois no Nordeste como um todo, se desenvolverá a cultura agroexportadora da cana-de-açúcar, do algodão e do tabaco (nessa ordem de importância). A cultura do fumo se desenvolverá predominantemente na Bahia, na região de Cachoeira, e desempenhará um importante papel, nos séculos XVIII e XIX, como elemento de troca no comércio de escravos. A cultura do algodão se difundirá no território dos atuais Estados do Ceará, Paraíba, e o alto sertão da Bahia e de Minas Gerais; mas sobretudo no Maranhão, que, apesar da forte presença indígena, receberá largos contingentes de escravos africanos, em razão da já referida resistência cultural indígena ao trabalho forçado e às campanhas movidas em seu favor pelo padre Antônio Vieira. Contudo, foi certamente a lavoura da cana-de-açúcar aquela que deu o tom dos primeiros séculos do período colonial (LUCCHESI, 2009b, p. 46).

Devido à convivência da população negra com a população indígena, nesse ambiente em que se desenvolveu a cultura agroexportadora, há possibilidade de ter acontecido o contato com a língua geral. No entanto, devido ao reduzido número de índios na produção açucareira, os africanos passaram a ter maior contato com o português. Desde o final do século XIX, mesmo que timidamente, já aparecem publicações sobre essa situação de contato linguístico e as influências das línguas indígenas e das línguas africanas na formação do português brasileiro, sendo que há um destaque maior para os estudos linguísticos afro-brasileiros. Um desses trabalhos é o de João Ribeiro através da publicação, em 1888, do *Dicionário Grammatical*, que teve uma nova edição “Inteiramente refundida e muito aumentada” em 1906. Além da influência das línguas indígenas, na formação do português brasileiro, Ribeiro (1906 [1888]) aborda, também, denominando de elemento negro, as alterações na língua brasileira, influenciadas pelas línguas africanas, a partir do povoamento da colônia com a vinda de um grande contingente de negros escravizados trazidos para o Brasil, em uma referência, na verdade, ao trabalho anterior de Macedo Soares (1880):

Temos muitas vezes, no correr deste escripto, falado em *dialecto brasileiro*. Cumpre observar que não apurámos o valor scientifico da palavra *dialecto*. Si entendermos por *dialecto* a linguagem derivada da lingua geral de uma nação e particular a uma cidade ou provincia, o brasileiro não é dialecto do portuguez. Por outro lado, a expressão *lingua brasileira* nos parece demasiado pretenciosa, si se quer com ella distinguir o portuguez falado no

Brazil, modificado pelo clima, pela natureza ambiente, pela influencia dos elementos africano e indiano, das relações commerciaes, etc, do portuguez falado em Portugal (MACEDO SOARES, 1880, p. 270-271).

Entre essas alterações, Ribeiro (1906 [1888]) abordava a questão da redução das formas verbais e também a simplificação das flexões de plural na fala popular do português brasileiro, enfatizando que essas alterações foram reforçadas pela influência indígena, já que encontramos, durante um longo período no Brasil, até o século XVIII, uma língua franca de base tupi-guarani, usada como forma de comunicação entre os portugueses e a população nativa e ainda entre os nativos de diversas etnias, conforme citado também por Lucchesi (2009b) acima.

As línguas faladas pelos africanos escravizados foram responsáveis por produzir alterações no vocabulário e no sistema gramatical da língua portuguesa, mas, além disso, há ainda a possível influência das línguas africanas sobre construções sintáticas, modos de dizer, além dos desvios populares do português da antiga metrópole. Sobre esse contato, um estudo de grande importância sobre as línguas africanas e o português do Brasil foi a publicação, em 1933, de *Os africanos no Brasil*, de Nina Rodrigues, que despertou um grande interesse pelos estudos afro-brasileiros. Rodrigues (2010 [1935]) propõe que sejam feitos estudos para saber quais as línguas africanas faladas no período colonial no Brasil, além de que sejam observados os modernos estudos feitos na África sobre elas para, em seguida, observar qual seria, então, a influência que teriam exercido no português falado no Brasil.

Também desse período, encontramos *O elemento afro-negro na língua portuguesa*, de Jacques Raymundo, em que há uma tentativa de sistematização dos fatos que teriam, de certa forma, constituído a língua dos africanos escravizados trazidos para o Brasil: “Uma linguagem própria, mesclada do idioma natal e do português, a que se juntou a contribuição vocabular do indígena, e que determinou as alterações ainda hoje notadas no fonetismo, no ritmo e na sintaxe de nossa fala popular” (RAYMUNDO, 1933, p. 75).

A maioria das propostas de Raymundo (1933) passa por um problema inerente às obras desse período, que é a limitação bibliográfica sobre os estudos das línguas faladas pelos africanos durante o período colonial em terras brasileiras no Brasil, o que ainda permanece até os dias de hoje. O que nos chama a atenção, o entanto, é a situação de contato relatada pelo autor sobre a língua dos africanos

no Brasil, que seria misturada com a contribuição vocabular do indígena e que determinaria alterações que influenciaram na fala do português popular do Brasil. Essa afirmação é de suma importância, já que vem mostrar o quanto essa questão sempre foi abordada e evidenciada. Não vamos adentrar na análise dessas palavras, pois o nosso interesse, neste caso, é mostrar os estudos que incluíram a influência das línguas africanas no português brasileiro e que esta não é uma questão somente abordada recentemente.

De qualquer forma, essa obra tenta classificar e analisar linguisticamente, pela primeira vez, as contribuições africanas para o português do Brasil, ou seja, palavras de origem africana que entraram para o vocabulário da língua portuguesa em território brasileiro. Nesse mesmo ano foi publicado *A influência africana no português do Brasil*, de Renato Mendonça, que teve uma segunda edição ampliada e ilustrada em 1935. Mendonça (2012 [1933]) busca, em uma parte, a sistematização referente ao português popular brasileiro relacionando o contato entre línguas africanas e o português europeu antigo. Esse estudo vai ser reproduzido com algumas alterações em outra obra sua, publicada em 1936, intitulada *O português do Brasil*.

Outros trabalhos ainda foram publicados a partir desse período, como em 1936, *Os africanismos no dialeto gaúcho*, de Dante de Laytano; *Africanos no Brasil*, em 1938, de Nelson de Senna; *A influência social do negro brasileiro*, em 1943, de João Dornas Filho; *O Negro e o garimpo em Minas Gerais*, em 1944 (com segunda edição em 1964), de Aires da Mata Machado Filho; *A linguagem popular da Bahia*, em 1951, de Édison Carneiro, entre outros. A maior parte dessas obras, no entanto, se limita a apresentar vocábulos de origem africana que se fazem presentes no português brasileiro, de origem banto, quimbundo, iorubá e assim por diante ou, ainda, termos africanos utilizados em religiões de origem africana cultuadas no Brasil ou associados a manifestações folclóricas brasileiras.

Também um estudo sobre o português no Brasil dentro dessa ordem cronológica de publicações e que aborda o mesmo tema das publicações citadas anteriormente é *A língua do Brasil*, de Gladstone Chaves de Melo, publicado em 1946. A questão da insuficiente e precária bibliografia produzida sobre esse assunto, que já citamos anteriormente, é abordada por Melo (1981 [1946]) em um dos capítulos intitulado “A influência africana”, afirmando ainda que a paixão pelo tema tem interferido no senso crítico dos pesquisadores. De fato, praticamente não há

registros da época, o que se deve inclusive à impossibilidade de serem feitos, já que em se tratando das línguas africanas no Brasil e do português popular brasileiro o que temos é a prevalência da oralidade, da língua falada, o que dificultaria essa tarefa.

Soma-se a isso a falta de interesse em uma mistura de línguas em um período em que o interesse seria impor a cultura europeia, mais precisamente a portuguesa. Algumas questões importantes abordadas por Melo (1981 [1946]) em sua obra dizem respeito à influência das línguas africanas no português brasileiro ocorrida na fala popular, que ele divide em vertical e horizontal. A primeira delas, a vertical, considerada mais profunda do que o tupi, ocorreu na morfologia, através do processo de simplificação e redução das flexões de plural e das formas verbais, sendo de grande importância para enriquecer a língua portuguesa.

A outra influência, denominada horizontal, seria a contribuição ao léxico, através das palavras, e, apesar de menos extensa que a tupi, teria também uma importância significativa em relação ao enriquecimento da língua portuguesa. Melo (1981 [1946]) admite a ocorrência de duas línguas gerais de origem africana, ambas na Bahia, o nagô ou iorubá, em Salvador, e o quimbundo em outras regiões, mas não é leva em consideração a existência do quicongo e do quimbundo, incluindo também a Bahia. Vale ressaltar essa referência à redução das flexões de plural e das formas verbais sempre que se fala da influência das línguas africanas na formação do português popular do Brasil, desde os trabalhos de Macedo Soares (1880) e Ribeiro (1906 [1888]), já citados acima e que parece prevalecer nos estudos sobre o português afro-brasileiro.

Contrariando essa visão da influência de línguas africanas ou indígenas, temos, por outro lado, a publicação, em 1951, com uma segunda edição em 1963, de *Introdução ao estudo da língua portuguesa no Brasil*, de Serafim da Silva Neto. No que se refere a essa questão da redução das formas verbais e das flexões do plural no português popular do Brasil, Silva Neto (1963 [1951]) associa às precárias condições sociais da população negra e indígena, o que levou a uma aprendizagem “tosca” da língua portuguesa. Esse estudo é baseado em elementos da etnografia e da história social do povo brasileiro, que evidencia a divisão do português brasileiro em diversos níveis, dependendo de ocasiões, regiões e classes sociais. De acordo com os estudos de Silva Neto (1963 [1951]), a influência africana se pôde notar tanto no meio urbano quanto rural, principalmente nas áreas onde houve uma

grande necessidade de mão-de-obra. Já a indígena, se restringe a uma influência rural, pois logo os índios deixaram as cidades.

No entanto, para Silva Neto (1963 [1951]) a diferenciação da língua portuguesa do Brasil da de Portugal, em se tratando da questão dialetal dos falares brasileiros, tem caráter arcaizante e conservador, sendo que a ação das línguas indígenas e africanas se deu, principalmente, no vocabulário, já que o português europeu nunca deixou de ser um modelo ideal em terras brasileiras. Entretanto, Silva Neto (1963 [1951]) admite a possibilidade de certas “interferências linguísticas”, sobretudo em se tratando do contexto linguístico e social do Brasil nos três primeiros séculos de colonização. De qualquer forma, vem à tona novamente a questão do contato entre línguas na formação do português brasileiro, o que não pode, de maneira nenhuma ser descartada, em se tratando da história linguística do Brasil. Dando continuidade ao pensamento de Silva Neto (1963 [1951]), Sílvio Elia (1979) publica *A unidade linguística do Brasil*, atribuindo o papel de superioridade da língua portuguesa em relação às línguas africanas e ameríndias. Vale destacar que a discussão sobre semicrioulos, ou crioulos, que provavelmente teriam surgido no Brasil colônia, voltam à discussão na referida obra.

Sobre a presença dos africanos no meio urbano e rural, citados anteriormente por Silva Neto (1963 [1951]), é importante também destacar que, dentre os escravos que falavam o português, conforme Lucchesi (2009b) havia a divisão entre *ladinos* e *boçais*. Enquanto os *ladinos* falavam e entendiam bem o português, os *boçais* não conseguiam se comunicar nessa língua. É possível também que tenha ocorrido uma língua franca africana entre os escravos usada para a comunicação entre eles. O fato é que um grande contingente de escravos africanos aprendeu o português de forma irregular, conforme Lucchesi (2009b), e de forma diferente entre aqueles que realizavam o trabalho doméstico e o trabalho rural. Esses utilizavam o português em contato com os europeus e faziam uso da língua materna nas senzalas. Voltaremos a falar sobre essas questões mais adiante.

Por outro lado, a expansão da língua portuguesa no Brasil se dá com o crescimento da produção da cana-de-açúcar e, posteriormente, com o ciclo do ouro, que favoreceu a vinda de grande número de portugueses para a colônia:

A grande onda migratória vinda de Portugal com o ciclo do ouro certamente favoreceu a difusão da língua portuguesa no Brasil, aumentando o acesso dos escravos aos modelos da língua-alvo do segmento dominante e

penetrando nas regiões do interior de São Paulo, Minas Gerais e Goiás, onde antes predominava a língua geral. Portanto, esses processos sociodemográficos podem explicar bem mais o retrocesso da língua geral no século XVIII, do que a tão propalada reforma pombalina e a expulsão dos jesuítas do Brasil, na década de 1750. O avanço da língua portuguesa no território brasileiro – seja em sua variedade nativa, veiculada pelos colonos brasileiros, seja na variedade defectiva, falada pelos escravos africanos e seus descendentes crioulos – dá-se primordialmente sobre uma base socioeconômica, com a expansão das lavouras de açúcar no século XVII e, sobretudo no século XVIII, com o grande ciclo das minas, cujo manancial extraordinário de riqueza teve um impacto sobre toda a economia mundial (LUCCHESI, 2009b, p. 48-49).

Outro fator bastante importante, em se tratando da formação da língua portuguesa no Brasil, diz respeito à vinda da família real de Portugal, fugindo, em 1808, das Campanhas Napoleônicas na Europa. Com a Corte Portuguesa, veio para o país um grande contingente de brancos portugueses e a elite política e cultural da época. Apesar da transferência da sede do império português para o Brasil, não houve mudança significativa na economia que permaneceu basicamente de atividade agroexportadoras, tendo como força de trabalho a mão de obra escrava. Somente em 1888, aconteceria a abolição da escravidão e, um após ano, a Proclamação da República.

Ao abordarmos essas questões sobre o português afro-brasileiro, que dizem respeito à socio-história, à linguística e à cronologia de trabalhos publicados no final do século XVIII até meados do século XX, percebe-se uma semelhança entre os fatores linguísticos abordados. O que nos chama atenção é a menção à ocorrência de simplificação morfológica resultando da redução das formas verbais e das flexões do plural no português popular do Brasil. Essa ocorrência é citada nos trabalhos de Macedo Soares (1880), Ribeiro (1906 [1888]), Melo (1981 [1946]) e Silva Neto (1963 [1951]). Ribeiro (1906 [1888]) cita essa influência africana, apesar de enfatizar que essas alterações são reforçadas pela influência indígena, uma língua franca usada no Brasil até o século XVIII.

De qualquer modo, fica clara a influência das línguas africanas, o que é reforçado por Melo (1981 [1946]), que considera essa mais profunda do que a tupi, inclusive citando a ocorrência de alterações na morfologia do português brasileiro, conforme já citado acima. Há ainda o fato da citação no trabalho deste autor das línguas gerais africanas na Bahia. Por outro lado, encontramos a visão Silva Neto (1963 [1951]), contrariando esses posicionamentos, ao afirmar que essas alterações ocorridas na simplificação das formas verbais e das flexões do plural no português

popular do Brasil foram resultado das precárias condições em que viviam as populações negras e indígenas, o que teria levado a uma aprendizagem ocorrida de forma irregular, malfeita, ou “tosca”, como ele mesmo denominou. Apesar de parecerem, a princípio, visões contraditórias, elas se complementam, pois a situação de contato linguístico do português europeu com as línguas indígenas e, posteriormente, com as línguas africanas em terras brasileiras, o aprendizado de maneira irregular da língua portuguesa, principalmente por parte dos africanos escravizados e de seus descendentes, tudo isso, provavelmente, contribuiu para a formação do português popular do Brasil.

Conforme Lucchesi (2009b), há um imenso campo de pesquisa a ser explorado quanto à compreensão da realidade da formação histórica da língua portuguesa no Brasil, já que há significativas diferenças entre o PB e o PE, o que tem sido, cada vez mais, objeto de pesquisa. Aspectos importantes da morfossintaxe (paradigma da flexão verbal, cliticização, parâmetro do sujeito nulo, movimento nas orações interrogativas) indicam mudanças afetando esses parâmetros à vista de valores não marcados, típicos de processos de criouliização (ROBERTS, 1997).

Impossível não falar sobre esse “caldeirão” linguístico e o contexto sócio-histórico que o envolveu, como também sobre a influência que toda essa situação teve na implantação da língua portuguesa no Brasil para se compreender as variedades linguísticas nacionais. Tanto no plano diatópico quanto no plano diastrático, para uma melhor compreensão das variedades linguísticas do PB e do PE, depende-se, e muito, de investigações apuradas sobre o processo histórico do contexto em que estiveram inseridas. Qual teria sido, então, a contribuição dos africanos na formação da sociedade brasileira? Qual o impacto e as consequências que ocorreram no campo linguístico?

Ao se falar do português afro-brasileiro, é importante salientar que esse se formou dentro de um contexto de contato linguístico que envolveu o português europeu, as línguas faladas pelos indígenas que já habitavam o Brasil e as diversas línguas vindas do continente africano com a população escravizada, que foi trazida à força para essas terras. Como este estudo refere-se especificamente ao português afro-brasileiro, nosso foco será a influência das línguas africanas na formação do português popular do Brasil. No entanto, não podemos deixar de abordar também a situação de contato com as línguas indígenas que eram faladas em terras brasileiras.

### 3.1 OS AFRICANOS NA FORMAÇÃO DA SOCIEDADE BRASILEIRA E O PLURILINGUISMO AFRICANO NO BRASIL

Ao se falar das línguas africanas que chegaram ao solo brasileiro, é necessário conhecer a realidade linguística do continente africano, que possui uma imensidão geográfica abrigando quase um terço das línguas do mundo. Essa grande quantidade, de 2152 línguas, segundo inventário publicado no *Ethnologue* (EBERHARD; SIMONS; FENNIG, 2022) pode ser ainda maior, com a ampliação de pesquisas posteriores, dividindo-se em quatro grandes troncos ou filios, de acordo com sistematização proposta por H. Greenberg nos anos 1950-1963: o nigerocongolês (Niger-congo), o afro-asiático (Afro-asiatic), o nilo-saariano (Nilo-saharian) e o coissan (Khoisan).

É difícil precisar, dentro da imensidão de línguas do continente africano, aquelas que estiveram envolvidas no ciclo de escravização e que chegaram até o Brasil no período colonial. Mesmo se apoiando em dados referentes ao tráfico negreiro durante aquele momento histórico, essa é uma pergunta muito difícil de ser respondida. O que se sabe é que a forma que os portugueses praticavam o tráfico e o comércio de negros escravizados favoreceu de certo modo, a seleção das línguas africanas, o que também contribuiu para a modificação linguística de algumas delas. Dentre os séculos XVI ao XIX, durante a vinda dos africanos escravizados para o Brasil eles conviveram, no mesmo período, com as línguas africanas faladas pelos africanos antes de atravessarem o Atlântico, bem como as línguas faladas pelos africanos, que já se encontravam escravizados em solo brasileiro. Viana Filho (1946) cita quatro ciclos grandes ciclos, entre outros, de importação de escravos na Bahia:

- (i) Ciclo da Guiné. Sec. XVI.
  - (ii) Ciclo de Angola. Sec. XVII.
  - (iii) Ciclo da Costa da Mina. Sec. XVIII.
  - (iv) Última fase. A ilegalidade. Sec. XIX.
- (VIANA FILHO, 1946, p. 28).

Vale salientar que essa divisão em ciclos não é totalmente estanque, já que a chegada de negros escravizados de uma região não impediu que continuassem a ser importados homens negros de outras localidades. Além disso, cada ciclo esteve associado a razões econômicas, sendo a cana de açúcar e do fumo, nos séculos XVI e XVII, a exploração de ouro e diamante, a produção de algodão, arroz e especiarias, no século XVIII, e a cultura do café, no século XIX. Em relação à

quantidade de africanos escravizados trazidos para o Brasil durante o período colonial, as informações são bastante divergentes. Não há fontes que comprovem esses dados. Lucchesi (2009b) cita que, normalmente, se leva em conta uma quantidade aproximada de três milhões e quinhentos mil africanos, o que é tomado como base quando se fala sobre a escravidão no Brasil:

Pandiá Calógeras (1927 [1957]) partiu de dados de recenseamentos da escravidão no início do século XIX, para chegar a uma média anual de cinquenta e cinco mil escravos importados. Desse modo, apresentou o número impressionante de treze milhões e quinhentos mil escravos importados, que a maioria dos estudiosos considera exagerado. Roberto Simonsen (1937), baseando-se na produtividade e na vida média do escravo, chegou a um número bem menor: três milhões e trezentos mil. Já Maurício Goulart (1949 [1975]), que adotou um método direto, com base nas poucas informações remanescentes sobre o tráfico, chegou a um número semelhante, que oscilava entre três milhões e quinhentos mil e três milhões e seiscentos mil (LUCCHESI, 2009b, p. 59).

Como se pode ver, há uma grande divergência no que se refere à quantidade de africanos trazidos para o Brasil no período do tráfico negreiro. Geralmente, se toma como base para os estudos sobre esse período a quantidade de três milhões e quinhentos a três milhões e seiscentos mil negros vindos da África. De qualquer forma, a imensa quantidade de africanos que vieram escravizados para o Brasil exerceu uma forte influência na constituição da sociedade brasileira, causando um grande impacto, com grandes e inevitáveis consequências também no campo linguístico. O recenseamento feito pelo Padre Anchieta em 1583, no início da importação de africanos escravizados, mostra que a população no final do século XVI era formada por 40% de brancos e 60% por negros e índios:

Os índios são mais representativos nas regiões laterais da Colônia, na época: São Vicente, Rio de Janeiro e Espírito Santo. Nos centros econômicos mais dinâmicos e progressistas – Pernambuco e Bahia – a presença dos portugueses, bem como de seus escravos africanos, é mais forte. Os colonizadores brancos constituíam a metade da população na Bahia, sede do Governo-Geral; e 40% em Pernambuco. Já os escravos africanos somavam quase 20% da população baiana, e constituíam a metade da população em Pernambuco (os números de escravos nascidos no Brasil, os crioulos, seria então irrelevante) (LUCCHESI, 2009b, p. 59-60).

No entanto, o número de africanos no Brasil provavelmente tenha sido maior no século XVI, pois há estimativas que indicam ter havido uma importação de escravos em torno de cem mil até o ano de 1600, aumentando ainda mais no século seguinte, quando vieram seiscentos mil escravos, conforme Taunay (1941, p. 256).

A partir do século XVIII, no entanto, a proporção de negros brasileiros e africanos passa a atingir a cifra de um milhão e trezentos mil, metade da população do Brasil. Esse teria sido o período mais favorável para que ocorressem processos crioulizantes com maior intensidade. Provavelmente essa proporção atingiu mais de 50%, Conforme Lucchesi (2009b), principalmente nas plantações de cana-de-açúcar do Nordeste.

A partir da imigração portuguesa durante os séculos XVIII e XIX, há um aumento de brancos e mestiços, sendo que a proporção de negros e africanos passa a ser constituída, em 1890, por apenas 15% do total. Desse modo, a quantidade da população branca em terras brasileiras nunca foi menor que 30%, segundo afirma Lucchesi (2009b, 2015) sendo que no século XIX chegou a um valor próximo da metade do total. Uma situação bem diferente do que teria ocorrido no Haiti e na Jamaica, por exemplo, em que a população de brancos em nenhum momento, no período da colonização, atingiu uma quantidade superior a 10%. Além disso, a miscigenação entre brancos e negros foi bem menor do que o ocorrido no Brasil.

Esses dados nos levam a questionar se realmente era homogeneamente 30% a 40% a quantidade da população branca em todo o território brasileiro naquele período. Mesmo que o cômputo total seja de 30% ou 40%, as relações eram homogêneas? Ou seja, o quantitativo total de brancos e pretos interagiu uniformemente em todo o território durante todo o período colonial? Essa questão é importante para se discutir quão relevante é a cifra realmente. Falaremos mais à frente sobre essa questão, ao apresentarmos as abordagens de Pagotto (2018). Lucchesi (2009b, p. 62-63) também afirma que “De um modo geral, o acesso aos modelos da língua-alvo no Brasil sempre foi maior do que o que se observa nas situações mais típicas de crioulização”.

Essa conjuntura vai se refletir no contexto linguístico, conforme afirma Lucchesi (2009b), que será tratado mais adiante:

O nível de crioulização é inversamente proporcional à facilidade de acesso aos modelos da língua-alvo. Em termos demográficos, o acesso à língua-alvo é determinado pela proporção da população dominante. Em outras palavras, quanto menor for a proporção do segmento dominante, maiores as chances de ocorrer a crioulização (LUCCHESI, 2009b, p. 62).

Essa situação não se aplica ao povoamento no interior do Brasil, como enfatiza Lucchesi (2009b) em nota de rodapé, já que em grandes propriedades rurais, o número de escravizados excedeu e muito em relação à proporção geral, e nos casos dos quilombos, onde se refugiavam os escravizados foragidos. Sobre essa questão das grandes propriedades rurais, Pagotto (2018) afirma que a vida nesses locais promovia uma tendência para o desenvolvimento de uma intensa e duradoura ebulição, o que não irá resultar necessariamente em uma sedentarização linguística<sup>37</sup>, sendo que, nesses casos, a dinâmica da variação se processa com mais intensidade.

Dessa forma, ao estudarmos a história da língua, através das formações resultantes de processos históricos, a relevância das situações de efervescência está no que elas são capazes de produzir em se tratando de sedentarização linguística: “Um exemplo clássico é o de comunidades remanescentes quilombolas que apresentam traços linguísticos dialetais diferentes do entorno geográfico em que estão inseridas” (PAGOTTO, 2018, p. 60). Nesses casos, podemos dizer que o estado de efervescência original, de certa forma se sedentarizou, passando a atuar, a partir daí, o que foi denominado por Pagotto (2018) como estabilização variacional sociolinguística, ou seja, a constituição de uma comunidade linguística, onde a língua passar a ser “reconhecida como índice identitário, passando a variar e a sofrer processos de mudança na dinâmica decorrente das relações sociais que se formam, dos papéis sociais constituídos em função de sua dinâmica econômica” (PAGOTTO, 2018, p. 60).

A partir desses conceitos, é possível pensar os diferentes cenários languageiros durante a história do Brasil, em que há uma diferença entre a estabilização sociolinguística nas cidades e nas regiões rurais. Na cidade, a dinâmica da variação se processa com mais intensidade, enquanto nas comunidades rurais, onde os papéis sociais são em menor número e com interações

---

<sup>37</sup> Por sedentarização linguística, entendo o fato de a língua se fixar em um dado território, perenizando-se por meio do processo de aquisição da língua. Esse processo se opõe a outro, que podemos chamar de ebulição ou efervescência linguística. Por ebulição ou efervescência linguística compreendo aquelas situações languageiras em que processos econômicos muito específicos constituem sociedades transitórias, em geral produto de migração de falantes de diferentes línguas ou dialetos. Tais situações podem dar origem a dinâmicas linguísticas muito aceleradas, mas também voláteis (PAGOTTO, 2018, p. 59).

sociais vinculadas a regras de comportamento mais duradouro, tende a ser mais moderada. Consequentemente, podemos conceber, desse modo, também a vida nas grandes propriedades com uma tendência para o desenvolvimento de uma intensa e duradoura ebulição, o que não irá resultar necessariamente em uma sedentarização no sentido aqui abordado por Pagotto (2018).

A formação do português brasileiro é o resultado de um imenso contato entre línguas, no entanto, Pagotto (2018) defende que isso seria uma consequência dos processos de sedentarização linguística, que também foram responsáveis, posteriormente, pela origem dos dialetos regionais:

A hipótese que trago é a seguinte: é nos contextos de subsistência econômica que vamos encontrar as condições fundamentais de sedentarização linguística que determinarão as feições do português brasileiro. Assim, não basta tomar o argumento demográfico cru, em si mesmo. É claro que a entrada de quatro milhões de africanos ao longo do período escravagista produziu infindáveis situações de efervescência linguística, com falares muito distantes do português e situações de multilinguismo incontáveis (PAGOTTO, 2018, p. 60-61).

Desse modo, iremos encontrar nesse grande número de homens e mulheres livres ou escravos, que movimentavam a economia do mercado interno, tanto nas cidades como nas regiões rurais, as origens do português brasileiro. Segundo Pagotto (2018), partindo desse ponto de vista, há possibilidade de se nivelar mais precisamente o papel desempenhado pelo contato do português com línguas e culturas indígenas e, dessa forma rever o papel que o contato linguístico teve na conformação do português brasileiro.

Partindo desse pressuposto, a partir do desenvolvimento das atividades de subsistência, se configura o processo de mudança da língua, sendo que, dessa forma, só encontraremos a ponderação adequada para o papel do contato na formação do português brasileiro, se levarmos em consideração como sendo de longa duração. Isso acarreta dificuldades preocupantes para outras análises que conjecturam uma ruptura inicial, atenuada posteriormente:

Deste modo, hipóteses como a da assim chamada Transmissão Linguística Irregular não seriam adequadas para dar conta da complexidade que os diversos quadros de contato linguístico trazem para a história do português brasileiro, porque tentam englobar, no seu interior, situações de emergência (decorrentes do que chamei de ebulição/efervescência) e situações de longa duração (como é o nosso caso, em que a sedentarização linguística

acompanha o processo de ocupação da terra por pequenas propriedades e a formação de pequenos núcleos populacionais) (PAGOTTO, 2018, p. 61).

Os modelos que têm como base a transmissão linguística irregular, de acordo com Pagotto (2018), subtendem uma ruptura inicial que, com o decorrer do tempo e a mudança das gerações, é atenuada pelo subsequente contato com a língua alvo dominadora, sendo a possibilidade de remanescentes crioulos em comunidades isoladas o argumento para esse tipo de hipótese. No entanto:

É preciso tomar cuidado com esse argumento de natureza teleológica, por mais sedutor que pareça: o fato de encontrarmos comunidades isoladas com traços linguísticos indicadores de uma ruptura maior não significa que temos aí preservado o que teria sido geral no restante do território. Pode ser justamente o oposto: o fato de estar em comunidades isoladas demonstraria apenas que houve condições específicas de sedentarização que permitiram a perenização de traços originários de uma situação específica de contato linguístico que, de resto, não teriam se repetido com tanta frequência como se acreditava no restante do território (PAGOTTO, 2018, p. 61-62).

Dessa forma, Pagotto (2018) retoma a afirmação de que é nas relações do mercado interno e o processo de autorreprodução deste que serão encontrados o vetor principal da economia que contribuiu para a constituição do português brasileiro:

É no interior das casas pobres de pau a pique e sem mobília, nos pequenos sítios, na multidão de carpinteiros, pedreiros, coureiros, vaqueiros, nas cidades dignas do nome, ou nas vilas pobremente envivecidas de gente em tempo de colheita, entre aqueles que podem criar e procriar, buscando na terra devoluta o fator primário de produção (para depois dela ser expulso, quando a expropriação pelo latifúndio a ela chega); é no seio das mães que aram ou que cuidam das pequenas vendas, das galinhas soltas em terreiros sem cerca, que vamos encontrar o lugar onde a língua deita seu corpo sobre a terra e se faz brasileira (PAGOTTO, 2018, p. 62).

Ao nos referirmos ao contexto da história linguística do Brasil, não podemos deixar de enfatizar esse aspecto da história afro-brasileira, citado acima por Pagotto (2018), Lucchesi (2009b) e também comentado por Mattos e Silva (2004) a seguir, como de importância fundamental, dentro de uma perspectiva histórica, tanto para a compreensão quanto para a difusão do português popular brasileiro: a formação dos quilombos durante o período colonial no Brasil, quando uma grande população de africanos escravizados foi trazida para o país e forçada a trabalhar, principalmente em grandes lavouras e, posteriormente, na exploração de minérios.

Essa questão se faz presente na bibliografia que se refere à formação da história do português brasileiro, conforme afirma Mattos e Silva (2004, p. 86), mas se reduz ao “megaquilombo” de Palmares, situado na atual região Nordeste do Brasil, que era um conjunto de mais de doze quilombos, que abrangia o que hoje é parte dos estados de Alagoas e Pernambuco, reunindo uma grande população. Para Schwartz (1987, 1994) há certo exagero ao se referir à população que vivia no Quilombo de Palmares, como em torno de 20 a 30 mil indivíduos, já que os engenhos de Pernambuco provavelmente nunca tinham tido esse número de escravos ao mesmo tempo:

Um relato de meados do século dezessete descreve Palmares como sendo dividida em dois agrupamentos principais e vários de tamanho menor, estimando a população dos vários agrupamentos em 11.000. Estimativa posterior e frequentemente repetida aumentava esse número para 30.000, o qual parece exagerado (SCHWARTZ, 1987, p. 81-82).

É importante salientar a existência de diversos grandes quilombos no século XVIII, de acordo com Schwartz (1987), o do Ambrósio, em Minas Gerais, e o Quilombo Grande, também mineiro, o quilombo do Piolho, no Maranhão, entre outros, do Rio Grande do Sul ao Amazonas, com uma grande concentração de população escrava, já que a geografia contribuía para o sucesso da fuga. Apesar da dificuldade em detectar a existência dos quilombos, somente na Bahia, segundo Schwartz (1987), foram identificados mais de 25, entre 1614 e 1825. Já em Minas Gerais existiram pelo menos 126 desses agrupamentos durante o século XVIII, sendo que os estudos históricos afro-brasileiros mais recentes têm revelado uma grande resistência à escravidão no período escravista, que teve um crescimento ainda maior no século XIX com diversas revoltas e rebeliões negras. Esses quilombos se espalharam pelo Brasil afora, como forma de resistência à escravidão e, ainda hoje, são encontrados remanescentes desses agrupamentos, com descendentes dessas populações africanas, que são uma importante fonte de contribuição para o estudo da formação do português brasileiro.

As pesquisas recentes sobre a historiografia afro-brasileira têm revelado que essa resistência foi constante ao longo de todo o período em que ocorreu a escravidão no Brasil, conforme Mattos e Silva (2004), e que houve uma progressão no século XIX, quando ocorreram diversas revoltas e rebeliões negras. Os quilombos se constituíam, então, de agrupamentos sociais que, apesar do

isolamento para a sobrevivência, também por esse mesmo motivo, necessitavam de uma articulação com a sociedade considerada legítima, como afirma Mattos e Silva (2004), o que deixa um forte indício em relação ao papel desempenhado pelos africanos e afrodescendentes na construção e difusão do português geral brasileiro.

Sendo assim, a forma de socialização da língua-alvo dos africanos e as condições de vida a que eram submetidas no Brasil são fatores que contribuíram para inibir a formação de um crioulo brasileiro, conforme abordaremos mais à frente.

A criouliização tem uma relação de dependência crucial da nativização da língua-alvo, que também passa por um processo de socialização do uso dessa língua-alvo entre a população dominada:

A socialização da língua dos dominantes entre a população dominada e a sua nativização colocam essa língua precariamente adquirida diante de necessidades expressivas e comunicativas novas e que não se apresentavam na simples interação imperativa com os indivíduos do grupo dominante. As lacunas do antigo código de emergência são então preenchidas através da gramaticalização, decalcadas ou não das estruturas originais das línguas do substrato, de formas lexicais da língua-alvo, retidas na aquisição do código de emergência. Esse processo de reestruturação é fundamental nos casos típicos de criouliização (LUCCHESI, 2009b, p. 63).

Vários outros fatores, no caso do Brasil, também prejudicaram uma socialização da língua-alvo entre os escravos, entre eles as precárias condições em que viviam o que dificultava a formação de uma vida social ou familiar. A exploração desumana a que eram submetidos levava a um alto índice de mortandade, com um ciclo de vida de trabalho entre cinco e seis anos. Desse modo, entre a população escrava havia uma grande dificuldade para constituir uma família ou procriar, apresentando uma taxa de crescimento vegetativo negativa. Também havia uma proporção maior de homens entre os que foram trazidos da África, sendo constatada uma proporção de três homens para uma mulher. Outro fator era a grande taxa de mortalidade infantil entre os escravos, segundo Mussa (1991, p. 155), em torno de 60%. Somam-se a tudo isso os deslocamentos consecutivos que ocorreram das lavouras do Nordeste para a região de exploração de minérios, durante o século XVIII, e depois para as lavouras do café no século XIX. Além disso, a venda de escravos, que era muito comum, desfazia a formação de grupos e até mesmo do início de um núcleo familiar, conforme afirma Lucchesi (2009b), dificultando assim a socialização entre eles.

Quando nos referimos acima que é praticamente impossível descartarmos a questão do contato linguístico entre línguas africanas e o português europeu, devemos também levar em conta a imensa densidade demográfica que, segundo Pessoa de Castro (2005), foi um dos fatores determinantes na presença de línguas africanas em território brasileiro durante o período colonial, com a presença estimada em cinco a oito milhões de africanos que foram trazidos para o Brasil com o intuito de substituir o trabalho escravo indígena.

Importante salientar que esses dados se diferenciam dos apresentados anteriormente por Lucchesi (2009b), mas estão, ao mesmo tempo, dentro de uma estimativa que deve ser considerada. Essa grande presença de africanos gerou uma população de 75% de negros e mestiços em relação à quantidade de portugueses e outros europeus presentes no Brasil, de acordo com o Censo de 1823, realizado um ano após a independência do Brasil. Esse fator, referente à superioridade demográfica, teve uma enorme contribuição, para o desaparecimento, a partir, inclusive, do século XVIII, do uso de uma língua franca de base indígena, conforme afirma Pessoa de Castro (2005, p. 63), o que favoreceu a “emergência de dialetos afro-brasileiros nas senzalas, plantações e minas, nos quilombos e, mais tardiamente, em núcleos urbanos da costa, até o seu estabelecimento em línguas litúrgicas”. Além do mais, foram originados, de acordo com Pessoa de Castro (2005), em comunidades negras isoladas, uma forma de falar especial, através de um vocabulário menor oriundo desse período.

Outro fator relevante que devemos considerar dentro do contexto histórico desse período, é o isolamento social e territorial da colônia em relação a Portugal até 1808, segundo Pessoa de Castro (2005), o que proporcionou um ambiente mais conservador,<sup>38</sup> aberto aos empréstimos das diferentes culturas, através de um interesse comum e uma tendência niveladora. É importante destacarmos a figura da “mãe preta”<sup>39</sup> convivendo no núcleo da família colonial, além do escravo ladino, já citado anteriormente, que por saber falar português desde cedo exerceu uma

---

<sup>38</sup> Esse fato se deveu ao monopólio do comércio externo brasileiro feito por Portugal até 1808, de acordo com Pessoa de Castro (2005), o que tornou o Brasil isolado do contato com outros países da Europa e até mesmo com Portugal, não recebendo influências das mudanças linguísticas ocorridas na metrópole durante esse período.

<sup>39</sup> “As mulheres negras que tinham seus bebês tornavam-se mães pretas ou amas de leite. Nestes casos estas mulheres eram obrigadas a deixar seus filhos morrerem ou sofrerem para serem a mãe preta da criança branca” (SILVA, 2012, p. 94).

influência em um número maior de ouvintes, já que tinha acesso a duas comunidades de falantes sociolinguisticamente diferenciadas. Já, a partir de 1808, com a abertura dos portos e a vinda da família real portuguesa para o Rio de Janeiro, houve a fixação de africanos escravizados, recém-trazidos da África, em um período em que a maioria da população já era formada por mestiços e crioulos, já nascidos no Brasil, muitos deles adaptados a costumes europeus. Por fim, com a extinção do tráfico da África para o Brasil, muitos africanos foram deslocados para diversas regiões brasileiras, a depender das atividades econômicas desenvolvidas e das necessidades de cada uma delas, o que ocasionou, conforme Pessoa de Castro (2005, p. 64), “uma presença constante em todas as regiões do território brasileiro sob regime escravista”.

Todos esses fatores contribuíram para inibir o surgimento de uma língua crioula em território brasileiro, desde a miscigenação até a integração do negro na sociedade do Brasil, já que grande parte dos escravizados tinha como destino o trabalho na lavoura o que, de acordo com Lucchesi (2009b), era um lugar onde a aquisição da língua portuguesa se dava de forma mais precária. Por outro lado, aqueles que eram destinados aos serviços urbanos e domésticos, provavelmente, tinham possibilidade de acesso a um modelo da língua-alvo mais próximo da variedade de seus senhores.

Além disso, deve ser considerado que a maior proficiência em português sempre trouxe mais vantagens aos escravizados, mesmo entre os que se encontravam nos engenhos (LUCCHESI, 2009b, p. 67). Isso porque aqueles que tinham um maior domínio da língua-alvo e se tornavam, portanto, bilíngues, transitando entre a casa-grande e a senzala, eram os preferidos para trabalhos domésticos, segundo Lucchesi (2009b), bem como para exercer funções de confiança nos engenhos, tanto como guardas pessoais de seus proprietários, como capitães-do-mato, conforme também afirma Pessoa de Castro (1990). Com o surgimento das minas, escravos são destinados para outras funções, diferentes daquelas exercidas por escravos rurais, que eram praticamente anônimos e viviam uma situação mais rígida, o que distanciava a casa-grande da senzala. Os escravos domésticos e urbanos e aqueles que foram para as regiões das minas podiam se destacar através de seu trabalho individual, o que lhe possibilitava comprar ou adquirir a própria liberdade.

Por outro lado, a forte presença dos africanos no Brasil contribuiu também para que ocorresse uma situação de plurilinguismo, devido à quantidade de línguas africanas faladas por essa população. O longo período em que o tráfico negreiro durou no Brasil também foi outro fator importante para que essa situação permanecesse durante muito tempo.

A última fase do tráfico negreiro se dá no século XIX, de acordo com Vianna Filho (1946). Algumas datas são muito significativas e vão marcar o fim de um período dessa prática que retirou forçadamente milhões de africanos de suas terras para trabalhar como escravos em um projeto de exploração de mão de obra em colônias de países europeus. O ano de 1815 marca o início dessa fase, com a proibição do tráfico negreiro ao norte do Equador. Em 1826, o Brasil e a Inglaterra assinam a convenção para a extinção do comércio de escravos na costa da África. O tráfico de escravos para o Brasil é proibido, através da lei de 07/11, em 1831. Em 1836, Portugal proíbe o transporte de escravos pelo mar. A partir de 1840, com a vigilância das águas do Atlântico pela marinha britânica e a campanha internacional contra a escravidão, aos poucos o transporte marítimo é neutralizado e finalizado, depois de três séculos em atividade. Em 1850, através da lei Euzébio de Queiroz, o tráfico de escravos no Brasil é extinto. A entrada por via terrestre de escravos em Angola, provenientes de Luanda, é vedada em 1854. No entanto, somente em 1878, a escravidão é oficialmente abolida (RANDLES, 1968). No Brasil, isso só viria a acontecer em 1888.

Todo esse cenário viria contribuir ainda mais para o plurilinguismo africano no Brasil. Diversos documentos testemunham essa situação, de acordo com escritos de Balbi (1826), ao comentar um encontro que teve com Maurice Rugendas, no Brasil, que ao interrogar muitos africanos escravizados, conseguiu bastante informações sobre as línguas dessa população. No entanto, Balbi (1826) lamenta ainda a perda de muitos documentos manuscritos que continham vocabulários molua, mina, caçanje, dentre outros, recolhidos por M. Rugendas. É importante salientar a presença dos termos mina e caçanje, designando esses como grupos linguísticos relevantes, em relação à quantidade de falantes. Desse modo:

Graças aos dados linguísticos recolhidos junto aos escravos africanos por Dias, no Brasil, no século XVII, foi elaborada a primeira gramática sistemática da língua *quimbundo*, a primeira também a mostrar a inadequação do modelo latino dos casos para interpretar as classes nominais. Graças ao mesmo tipo de dados, mas desta vez coletados no

século XIX, na aurora mesma do estabelecimento da linguística comparada pelos irmãos Schlegel (1808 e 1818), uma miniexperiência comparativa permitia a A. Balbi classificar três línguas africanas do tipo *banto*, alguns decênios antes dos grandes trabalhos de W. H. I. Bleek. É por um método análogo que, vinte e cinco anos mais tarde, S. Koelle (1854) escreverá *Polyglotta africana* com base em dados recolhidos junto a antigos escravos em Freetown, em Serra Leoa, no qual ele estabelecerá uma classificação geotipológica de cerca de 300 línguas africanas (BONVINI, 2014, p.46-47).

No Brasil, o plurilinguismo africano é atestado de modo efetivo a partir do final do século XIX, o que foi testemunhado em 1890-1905<sup>40</sup> por Nina Rodrigues, em resposta a um apelo feito por Sílvio Romero, em 1879-1880<sup>41</sup>, que alertou sobre a necessidade de estudos urgentes sobre as línguas e as religiões africanas, sendo citado textualmente por Rodrigues (1935, p. 11-12): “Apressem-se os especialistas, visto que os pobres moçambiques, benguelas, monjoos, congos cabindas, caçanjes... vão morrendo. O melhor ensejo, pode-se dizer, está passando com a benéfica extinção do tráfico”. Foi, então, dessa maneira, que Rodrigues (1935) realizou, com os africanos que viviam na cidade de Salvador, sua pesquisa linguística e etnográfica naquela época.

O levantamento feito por Rodrigues (1935) foi bastante relevante, já que a partir da segunda metade do século XIX instalou-se no Brasil uma nova modalidade de tráfico, com o desenvolvimento da cultura do café, de acordo com Vianna Filho (1946), deslocando o eixo econômico do país para o Sul, o que, conseqüentemente, provocou também o deslocamento da população negra escravizada para o trabalho agrícola naquela região. Desse modo houve uma desconcentração econômica nesse período em que “A Bahia despovoava-se de escravos. Em 1815 teria 500.000. Em 1874 não seriam mais de 173.639” (VIANA FILHO, 1946, p. 90).

Rodrigues (1935) realizou um levantamento, na época, com uma lista contendo 122 palavras, que pertenciam a cinco línguas africanas faladas durante aquele período, de maneira corrente, na Bahia. São elas: “nagô”, “gege”, “hauçá” “kanuri”, “tapa” ou “nifê”, “grunçe”. Divide, então, as línguas africanas faladas no Brasil em: línguas do norte, línguas sudanesas, ou do centro, e línguas austrais ou banto (RODRIGUES, 1935, p. 195-197). Além disso, precisa que, entre as línguas africanas, “Dentre ellas, duas, as que foram adoptadas como línguas geraes,

<sup>40</sup> Para os nossos estudos, utilizaremos a edição de 1935 da obra de Nina Rodrigues.

<sup>41</sup> Sílvio Romero, “Estudos sobre a Poesia Popular do Brasil”, Rio, 1888, p. 10-11. Publicado após o Prefácio de Africanos no Brasil, por Nina Rodrigues (1935).

predominavam no paiz; a nagô ou iorubá na Bahia, o kimbunda ou conguesa no norte e no sul” (RODRIGUES, 1935 p. 197). Uma observação importante é o fato das línguas africanas faladas no Brasil terem sofrido profundas alterações de imediato, tanto a partir da aprendizagem do português por parte dos negros escravizados, quanto através da língua geral africana adotada pelos negros chamados de ladinos<sup>42</sup>:

Antes, no número das importadas, na infinita multiplicidade e matizes dos seus dialectos, ellas eram tantas que, num exaggero quasi desculpavel, se poderiam dizer equivalentes em numero ao dos carregamentos de escravos lançados no paiz. Em taes condições, tornou-se uma necessidade imperiosa para os escravos negros adoptar uma língua africana como língua geral, em que todos se entendessem (RODRIGUES, 1935, p. 189)

O nagô, ou iorubá, já citados anteriormente, conhecido como língua da Costa, segundo Rodrigues (1935), foi considerado como uma das línguas gerais mais importantes desse período. Falada, principalmente, por africanos mais velhos e por crioulos e mulatos, em grande quantidade, segundo Rodrigues (1935, p. 201), que afirmou: “Ela possui mesmo entre nós uma certa feição literária”, sendo, na África Ocidental, uma das primeiras línguas a conhecer uma tradição escrita.

O que leva a uma resistência para levantar questões relativas às línguas africanas no Brasil parte, principalmente, pelo prestígio que se dá à escrita, que é supervalorizada em relação à oralidade, de acordo com Pessoa de Castro (2005, p. 65), “a partir de uma pedagogia, vigente no mundo ocidental, que sempre privilegiou o ler e o escrever diante da não menos importante e mais antiga arte do falar e do ouvir”. Por isso mesmo, os povos que conhecem alguma forma de escrita literária são considerados superiores àqueles que se utilizam de uma tradição oral. Dessa forma, por exemplo, haveria uma “suposta inferioridade cultural do povo banto face aos oeste-africanos ou “sudaneses” no Brasil” (PESSOA DE CASTRO, 2005, p. 65).

Desse modo, quanto às línguas africanas que eram faladas no Brasil no período colonial, poucos são os documentos que registraram esse acontecimento linguístico. O primeiro deles data do fim do século XVII, escrito no ano de 1697 pelo sacerdote jesuíta Pedro Dias, no Brasil, e publicado em Portugal: *Arte da língua de Angola*. Esse texto é uma comprovação histórica de que as línguas africanas foram

---

<sup>42</sup> Assim se chamavam no tempo do tráfico os escravos negros que já conheciam a língua, os usos e costumes do país (RODRIGUES, 1935).

utilizadas no Brasil naquele período colonial, segundo Bonvini (2014), mostrando como era estabelecida a política linguística pelos jesuítas. A sua importância linguística é muito relevante por ser a primeira descrição gramatical do quimbundo, com dados dessa língua falada no Brasil, mostrando o olhar de um falante do português acerca de uma língua africana daquela época, além de ser um registro das línguas faladas no Brasil naquele período.

A *Arte da língua de Angola* do Padre Pedro Dias, é registro do emprego cotidiano e habitual durante o século XVII, de uma língua africana no Brasil falada por negros escravizados originários de Angola, o quimbundo, numa extensa área, que não se limita somente à Bahia. Por outro lado, de acordo com Bonvini (2014) também demonstra o olhar de um falante do português culto acerca de uma língua africana, com certa estranheza e espanto, já que este tem como padrões de referência tanto o latim como o português

Outro registro também importante sobre as línguas africanas no Brasil foi escrito em Minas Gerais por Antonio da Costa Peixoto sobre uma língua “mina”. O texto foi escrito 34 anos depois da publicação de *Arte da língua de Angola*, de Pedro Dias. Uma primeira versão foi escrita em 1731 e uma segunda, em 1741, em que foi acrescentado um prólogo, em duas páginas, e uma “advertência” final. Esse texto, conforme Bonvini (2014) possui sua relevância por se tratar de uma experiência inédita, já que ocorre em um espaço limitado, o quadrilátero mineiro (Vila Rica, Vila do Carmo, Sabará, Rio dos Montes), e de uma necessidade específica, que era explorar diamantes encontrados durante o século XVIII.

A população atingiu uma média de 100.000 africanos escravizados por ano. “Eles eram originários da costa do Benim – chamada Mina e situada *grosso modo*, entre Gana e a Nigéria – e renovados regularmente durante um período de uns 40 a 50 anos” (BONVINI, 2014, p. 39). A língua mina seria um falar veicular, segundo hipótese de anterior de Bonvini e Petter (1998), que se encontrava em uma fase de pidginização, de acordo com Manessy (1995, p. 22-27), sendo submetida: à adaptação, que é a reinterpretação, de acordo com um modelo estrangeiro; à simplificação, que é a diminuição dos mecanismos gramaticais e melhorias de seu rendimento funcional; e à redução total de seu sistema linguístico (Cf. BONVINI, 2014, p. 45). A língua mina, então, falada pelos africanos que foram deslocados para o trabalho de exploração de ouro e diamante, a partir do século XVIII no quadrilátero mineiro, seria de predisposição supralocal, originária de línguas africanas do grupo

*gbe* (subfamília *cuá*), criada no Brasil para facilitar a comunicação entre vários falares tipologicamente semelhantes, passando a ser constituída pela praticidade linguística e como uma solução mais econômica.

O uso de línguas francas africanas entre os escravos, que ocorreu com mais frequência nas senzalas, teria também contribuído para inibir a formação de um crioulo brasileiro:

A criouliização depende do preenchimento de lacunas que se dá quando a língua dominante é também usada na interação entre os escravos. Se, nessa interação, os escravos lançam mão de uma língua franca baseada na língua nativa do segmento étnico africano mais representativo e/ou de maior prestígio, o processo de criouliização é inibido. O uso de línguas francas africanas no Brasil não está totalmente esclarecido, e o tratamento dessa questão conduz a uma ligeira digressão sobre a composição etnolinguística dos escravos africanos trazidos para o Brasil (LUCCHESI, 2009b, p. 63-64).

Essa divisão é feita baseada em dois grandes grupos linguísticos: o *kwa* e o *banto*, havendo a predominância no grupo banto, entre os escravos trazidos para o Brasil, de línguas étnicas majoritárias, como o quimbundo, o quicongo e o umbundo. Enquanto no grupo sudanês “os seus principais representantes no Brasil foram os nagôs ou iorubás, e os jejes ou povos de língua *ewe*” (PESSOA DE CASTRO, 1990, p. 98-99).

Sobre essa questão etnolinguística e a origem dos africanos trazidos como escravos para o Brasil, de acordo com Mendonça (2012 [1933]), pode ser dividida em dois grandes grupos linguisticamente falando: o *kwa*, conhecidos como sudaneses, que eram da região do Oeste-Africano, aproximadamente, ao norte da linha do equador; e o banto, que estava situado na extensão de terras ao sul do equador. Segundo Mendonça, 2012 [1933], p.16: “Os sudaneses apresentam uma grande fragmentação linguística oposta à unidade substancial das línguas banto”.

Apesar da composição etnolinguística do Brasil ter se alterado no decorrer dos séculos, durante o período do tráfico de escravos em terras brasileiras, há uma predominância de escravos que foram trazidos da zona linguística banto, conforme afirma Lucchesi (2009b). No entanto, essa situação é modificada com a rota que liga a Costa da Mina à Bahia, o que favoreceu a vinda de falantes de língua *kwa*, principalmente do iorubá. A influência africana no Brasil, de acordo com a grande maioria dos estudiosos, é dividida em banto (Rio de Janeiro e todo o Sudeste e Pernambuco para o Norte) e iorubá (Bahia). O predomínio do banto ocorreu

principalmente nos séculos XVI e XVII, o que teria refletido na formação de línguas francas africanas no Brasil. As línguas quicongo, quimbundo e umbundo predominaram nas senzalas, de acordo com Pessoa de Castro (1990), tendo maior prestígio sociológico, o que pode ter favorecido a formação de línguas gerais africanas no Brasil. Desse modo, a comunicação entre os escravos isolados em condições precárias nas senzalas e entre os que estavam foragidos nos quilombos, em muitos locais, provavelmente tenha sido uma língua geral de base banto, indiferente ao português adquirido de forma precária para a forçada convivência com os portugueses, que eram seus senhores.

Outras línguas francas também foram utilizadas em diferentes regiões do Brasil, oriundas de outros grupos linguísticos africanos. Uma delas teve como base a língua *fon*, através de escravos falantes de língua do grupo *fongbe*, que foram importados pela Bahia e depois enviados para o trabalho na região das minas “que foi atestada na primeira metade do século XVIII, na região de Vila Rica” (PESSOA DE CASTRO, 1990, p. 107).

Sobre essa questão, reafirma Lucchesi (2009b, p. 66): “Essa língua franca veicular fon deve ter convivido com outras línguas francas de base quimbundo que provavelmente eram usadas entre os escravos introduzidos no Porto do Rio de Janeiro”. Devido a grande quantidade de falantes de línguas banto no Sudeste, essas línguas francas de base quimbundo caíram em desuso com o tempo, sendo substituídas por variedades da língua portuguesa que foram reestruturadas pelos descendentes dos escravos que passaram a viver nesses locais. Essas línguas se conservaram em comunidades rurais afrodescendentes.

De acordo com Lucchesi (2009b), o uso dessas línguas francas africanas teve sua primeira evidência com a descoberta de uma língua veicular de base lexical banto, por Aires da Mata Machado Filho, em 1944, na localidade de São João da Chapada, no Norte de Minas Gerais. Essas línguas teriam sido conservadas, como línguas secretas ou como afirmação de identidade étnica, segundo Lucchesi (2009b), sendo ainda presentes, atualmente, em comunidades rurais negras. Como exemplos, temos a falange descoberta na comunidade de Cafundó, em São Paulo e da língua do negro da costa, em Tabatinga, Minas Gerais, em que foram encontradas em ambas um léxico de base banto, principalmente quimbundo, com as estruturas gramaticais do português popular do Brasil (Cf. LUCCHESI, 2009b, p. 66-67).

O mesmo ocorreu na Bahia, onde se concentraram majoritariamente escravos falantes da língua kwa, de maioria iorubá, o que tornou favorável também para a utilização de uma língua franca iorubá, também chamada de *nagô*, que era muito frequentemente usada na Cidade da Bahia, atual Salvador, no século XIX, conforme citado já anteriormente por Rodrigues (1935). Seu uso provavelmente se prolongou até o século XX, conforme afirma Lucchesi (2009b), sendo que veio a se tornar uma língua ritual nos candomblés da Bahia.

Outro fato demonstrado historicamente em relação às línguas africanas no Brasil é de que elas, em sua essência, foram línguas orais. Havia escrita apenas no quimbundo, de acordo com Pessoa de Castro (2005), e em uma frequência menor na língua “mina”. No entanto, o tempo eliminou aos poucos essas ocorrências, já que elas foram produzidas por falantes aloglotas, o que, de qualquer forma, contribuiu para esse apagamento. Mesmo estando ausentes a escrita e a gramatização, as línguas africanas possuíam uma codificação que se relacionava com a fala, manifestando-se através de ‘interditos’ de enunciados referentes ao espaço, ao tempo e aos interlocutores e ainda, segundo Bonvini (2014, p. 57), “sob forma de aprendizagem das manipulações das imagens e dos símbolos (análoga a nossa retórica), conjunto de regras específicas das sociedades ditas de tradição oral”. Formas semelhantes da língua e da fala são encontradas no Brasil, que estão presentes em cultos como o candomblé, em que ainda existe uma grande tradição oral de origem africana, de acordo com Pessoa de Castro (2005) e Lucchesi (2009b). Isso ocorre, devido ao valor próprio desse culto conferido à noção de fala, que é ritualizado através da iniciação e na adivinhação, e devido aos textos orais utilizados, como narrativas, contos, provérbios, cânticos e outros ligados às tradições e atividades rituais. O mesmo não acontece em relação à umbanda, em que predomina o português.

O que abordamos acima nos remete, em grande parte, ao contexto encontrado no local em que realizamos a gravação dos nossos inquéritos. Quanto às tradições mantidas por essas comunidades remanescentes de quilombos, podemos perceber na comunidade de Lagoinha, conforme citaremos mais adiante, ao abordarmos a descrição da comunidade estudada, a manutenção de muitas dessas tradições, como a oralidade na divulgação de narrativas diversas, manifestações culturais, como samba de roda, samba de bumba, bumba meu boi, terno de reis, pila do café, produção de artesanato, entre outras.

Falaremos a seguir sobre a relação da língua portuguesa falada no Brasil e a influência das línguas africanas faladas pelos africanos escravizados, durante o período colonial, na formação do português brasileiro.

### 3.2 A LÍNGUA PORTUGUESA FALADA PELOS AFRICANOS ESCRAVIZADOS NO BRASIL E A INFLUÊNCIA DAS LÍNGUAS AFRICANAS NA FORMAÇÃO DO PORTUGUÊS DO BRASIL

É evidente que os africanos trazidos como escravos para o Brasil desempenharam, em relação à realidade linguística no país, um relevante papel na sua constituição. O contexto histórico mostra que, em relação à população indígena que aqui habitava antes da chegada dos portugueses, o grande contingente de africanos escravizados, que veio para o Brasil no período colonial, teve um papel mais relevante em se tratando de contato entre línguas na nossa história sociolinguística. Os indígenas sentiram, a partir dos primeiros contatos, o impacto de uma nova cultura imposta pelos portugueses e, segundo Lucchesi (2009b), ao longo dos primeiros cinco séculos foi promovido um grande genocídio desses povos, tantos pelas expedições expulsá-los de suas terras, capturá-los ou por causa de moléstias que não haviam adquirido nenhuma imunidade. Qualquer simples virose era capaz de dizimar tribos inteiras. Quando eram expulsos, saíam em direção às regiões mais inacessíveis, buscando refúgio, fugindo da verdadeira chacina praticada pelos portugueses. Ao serem escravizados, não conseguiam se adaptar aos trabalhos forçados, assim como à vida urbana, em função das características culturais peculiares ao seu povo. Sendo assim, a mão de obra utilizada na colônia brasileira foi a do trabalhador escravo, através dos africanos trazidos à força para o projeto agroexportador de Portugal. Além disso, eles foram utilizados também como trabalhadores domésticos e para serviços urbanos, conforme afirma Lucchesi (2009b, p. 58): “Chegaram mesmo a desempenhar as funções de controle da produção – como capatazes – ou mesmo de repressão – eram negros muitos capitães do mato que caçavam escravos foragidos”. Sobre essa questão dos africanos que organizaram fugas e agrupamentos de resistência à escravidão, falaremos mais adiante.

Ao abordarmos a presença dos africanos no Brasil, não podemos deixar de mencionar a forma como eles chegaram até o nosso continente a partir da África. O comércio de negros africanos escravizados era feito a partir de locais de embarque

que passaram a ser cada vez mais fixos e a ter uma maior organização com o passar do tempo e com a experiência adquirida pelos portugueses. Poderia ser através dos barracões construídos nesses locais, com empregados para reunir em grande quantidade os cativos e depois entregá-los aos negreiros, ou ainda em navios ancorados em um porto próximo. Foram várias as línguas vindas do continente africano, mas algumas situações foram responsáveis pelo não estabelecimento dessas línguas no Brasil, sendo que há um consenso entre diversos autores, conforme Mattos e Silva (2004), no que se refere a duas das principais razões para que isso ocorresse.

A primeira diz respeito à forma como os africanos eram separados ainda na África para que não tivessem entre si relações étnicas ou linguísticas, uma das formas encontradas para que não acontecessem rebeliões contra o sistema escravista. A segunda está relacionada à impossibilidade de se constituírem laços familiares entre os escravos, conforme já citado anteriormente, o que impediria também a formação de núcleos linguísticos africanos. No entanto, durante o período em que permaneciam ainda na África, como prisioneiros, ou até mesmo durante o período de viagem até o Brasil, esse contexto favoreceu a emergência de uma situação linguística própria dos cativos que se encontravam nesses locais, principalmente com a continuidade do tráfico negreiro e com o objetivo de suprir o mercado brasileiro.

Uma situação transitória, devido à duração e sempre renovação dos negros aprisionados, que ficavam nos barracões ou nos navios por um determinado período, antes de serem enviados para o Brasil, onde esse contato linguístico ocorreria por um período maior. De qualquer forma, essa concentração forçada de falantes de diversas línguas africanas, que apesar de diferentes, por outro lado, eram próximas tipologicamente, conforme Bonvini (2014) levou, por exemplo, à adoção do *quimbundo* por Angola como língua veicular, sendo que era falada em Luanda e ao longo do rio Cuanza até sua foz.

Durante esse período, já havia variantes locais integradas à língua portuguesa falada em Angola, com uma grande diversidade de termos africanos, o que provavelmente já deveria ocorrer também no Brasil. Isso é comprovado em registros dessa época, nos testemunhos de escritos, como a *Colectânea documental de Fernão de Sousa* (1622-1635), com 87 termos; Oliveira de Cadornega (1680), com 160 termos; e a edição italiana da narrativa de Cavazzi (1687), com mais de

400 termos (Cf. BONVINI, 2014, p. 32). Ao mesmo tempo, na ilha de São Tomé se formava um crioulo, o qual se tornou a língua oficial daquele país, devido ao contato prolongado entre os negros escravizados que circulavam no local. Ali nasceu um *pidgin*, que era falado pelos descendentes de antigos escravos, que vieram de Angola em um navio para as Américas, ficando encalhado na ilha por causa de um naufrágio.

Nesse mesmo período, havia o contato, também forçado e prolongado, com a língua portuguesa, através da fala dos pombeiros, no interior, que eram os recrutadores dos africanos escravizados. Esse contato com o português se prolongava com os negreiros brasileiros, na costa africana, que podiam ser seus futuros senhores, além do período que se passava na viagem marítima que durava dois meses, trinta e cinco ou quarenta dias, a depender do local do Brasil para onde a embarcação se deslocava.

Na chegada ao Brasil à comunicação poderia se dar de diversas maneiras, através dos recursos linguísticos disponíveis, como as línguas indígenas, as línguas gerais indígenas e o português do colonizador, sendo que essa última opção foi a selecionada. Essa questão fica evidenciada, segundo Mattos e Silva (2004), conforme indica o desenvolvimento histórico do Brasil. Esse português do colonizador, adquirido pelos africanos em um contexto de oralidade, foi reestruturado profundamente, sendo que além de ter um uso recorrente em toda a colônia, alguns desses negros escravizados já haviam tido acesso a alguma forma do português, possivelmente adquirido durante a rota do tráfico. Apesar da existência de áreas de concentração negra e negro-mestiça e de forte presença indígena no Brasil amazônico, de acordo com Mattos e Silva (2004), há fortes evidências da generalizada presença dos africanos e afrodescendentes em praticamente toda a extensão territorial do Brasil colonial.

O fato é que, ao chegarem no Brasil, os negros escravizados tiveram contato com outra realidade linguística e cultural. Houve, na verdade, uma grande ruptura com todo o contexto em que viviam no continente africano. O contato linguístico com novas línguas, conforme já citado acima, foi de imediato o que promoveu noções diferentes e novas denominações, já que conviviam no Brasil colonial, línguas indígenas, línguas africanas e a língua portuguesa falada pelos senhores. O negro escravizado era obrigado a falar o português, conforme Nina Rodrigues (1935), no momento em que desembarcava em terras brasileiras. Essa era uma exigência para

que pudesse haver certa comunicação e entendimento, daqueles que chegavam do continente africano, com os senhores brancos, com os mestiços e negros crioulos que habitavam aqui no Brasil:

Desta arte, ao desembarcar no Brasil, o negro *novo* era obrigado a aprender o português, para falar com os senhores brancos, com os mestiços e os negros creoulos, e a língua geral para se entender com os parceiros ou companheiros de escravidão (RODRIGUES, 1935 p. 189).

Apesar da presença do português europeu, desde a chegada em terras brasileiras, em 1500, até os séculos seguintes, perfazendo um total de 30% durante os três séculos de colonização, de acordo com Mattos e Silva (2004), conforme já citado acima por Lucchesi (2009b), e das línguas gerais indígenas aqui formadas, deve-se destacar, também, a população africana e os afrodescendentes ao se falar em multilinguismo e multidialetismo durante o Brasil colonial.

Sendo assim, Mattos e Silva (2004) chama atenção para o que seria um português geral brasileiro, que teria como principais difusores a significativa parcela de africanos e afro-brasileiros no período colonial, e que teria antecedido, historicamente falando, o português popular do Brasil. Darcy Ribeiro (1995) enfatiza a contribuição do negro como principal agente difusor da língua do colonizador, que aprendeu ouvindo os berros do capataz e a refez tornando o português brasileiro singular, possibilitando a sua difusão em todo o território brasileiro.

A presença dos africanos escravizados na sociedade colonial brasileira foi grande nos projetos de exploração mercantilista colonial. No entanto, as funções desempenhadas pelos negros, de um modo geral, incluindo aí os que eram livres e seus descendentes, vão, além disso, já que eles também desempenhavam diversos papéis, não só no interior das famílias dos colonizadores, mas também exercendo atividades externas nos núcleos urbanos que se encontravam em formação e no ambiente rural na colônia. Essas afirmações de Mattos e Silva (2004) corroboram a tese da importância da presença africana quando se fala no português popular brasileiro atual.

Certamente, essa população circulava nos lugares onde se encontrava esse português geral brasileiro, conforme Mattos e Silva (2004), que leva à pressuposição de que seria bastante diversificado e variável, devido às condições sócio-históricas que contribuíram para que isso viesse a ocorrer, apresentando desde variedades pidginizadas a variedades mais próximas do português europeu, havendo até a

possibilidade da ocorrência de línguas crioulas: “Tudo a depender da história social e sociolinguística dos indivíduos, de grupos de indivíduos e, no caso de línguas crioulas, a depender da sócio-história linguística das comunidades” (MATTOS E SILVA, 2004, p. 85).

Há ainda outra situação que está relacionada a essa população de africanos que veio trazida forçada para o Brasil, a dos negros escravizados foragidos, citados por Lucchesi (2009b) anteriormente, que se refugiavam em agrupamentos denominados quilombos, que, através da sua constituição humana e social, são de fundamental importância para constituição da história linguística do Brasil:

Eles eram agrupamentos sociais que, isolados para sobreviver, também, por causa exatamente de sua sobrevivência, se articulavam com a sociedade legítima, o que traz um indício interessante sobre o papel desempenhado pelos africanos e afrodescendentes na construção e difusão do português geral brasileiro (MATTOS E SILVA, 2004, p. 87).

Desse modo, por meio do contato entre os diversos habitantes desses quilombos, incluindo negros, índios e brancos pobres, em termos linguísticos, circulavam também as diversas falas encontradas no Brasil da época, de acordo com Mattos e Silva (2004), sendo elas africanas, indígenas, português africanizado, português indígena e ainda o português europeu. Isso se deu devido ao fato de que muitos fugitivos se abrigavam e eram aceitos nesses espaços. Por isso mesmo:

A constituição humana e social dos quilombos apresenta significativo interesse para a história linguística do Brasil. Eles eram agrupamentos sociais que, isolados para sobreviver, também, por causa exatamente de sua sobrevivência, se articulavam com a sociedade legítima, o que traz um indício interessante sobre o papel desempenhado pelos africanos e afrodescendentes na construção e difusão do português geral brasileiro (MATTOS E SILVA, 2004, p. 87).

Dentro desse contexto, é importante destacar os principais responsáveis em relação ao contexto de multilinguismo e multidialetismo no Brasil, durante o período colonial:

- a. o português europeu na sua dialeção diatópica, diastrática e diacrônica, que teria ao longo do período colonial um contingente médio 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 XVIII;
- b. as línguas gerais indígenas, que, plurais e dialetizadas, poderiam até confundir-se com o português geral brasileiro 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 da população africana e dos afrodescendentes que perfizeram uma classe média de mais de 60% da população por todo o período colonial (Mattos e Silva, 2004, p. 90).

O que é importante salientar, no entanto, é a forma como esse português foi adquirido pelos africanos e como eles passaram a utilizar essa segunda língua, imposta pelo europeu, que tinha como intuito somente a exploração da mão de obra escrava, conforme veremos mais adiante. Desse modo, é importante observar como se configurou o panorama linguístico brasileiro contemporâneo para entendermos o contato entre línguas no Brasil e como este se processou no decorrer da história da sociedade brasileira, sendo que algumas questões sócio-históricas e demográficas tiveram uma importância fundamental:

- (i) número de escravos africanos trazidos para o Brasil e impactos, em termos demográficos, desse contingente sobre a composição étnica da sociedade brasileira;
- (ii) origem geográfica e caracterização etnolinguística dos africanos e sua distribuição pelo território brasileiro;
- (iii) formas de interação desses indivíduos, nos níveis social, cultural e linguístico, com o grupo dominante (LUCCHESI, 2009b, p. 58).

Ainda está por ser feito um estudo mais sistemático e real de como os africanos se exprimiam usando a língua portuguesa no Brasil do século XIX, período marcado por dois acontecimentos importantes: a abolição da escravatura, em 1888, e, a seguir, o período da cultura do café, que modificou a distribuição econômica brasileira. Esse fato promoveu o deslocamento dos ex-escravos e de seus descendentes para outras regiões do Brasil, redistribuindo assim essa população concentrada em determinadas áreas geográficas, conforme já citado anteriormente por Mussa (1991), Lucchesi (2009b) e Pessoa de Castro (2005).

De imediato, a abolição da escravatura contribuiu progressivamente para o declínio e extinção das línguas africanas, que possuíam uma forte ligação com a escravidão. Essas línguas, durante três séculos foram sustentadas pela presença de uma grande quantidade de escravos em zonas bem definidas, geograficamente falando, em função da produção econômica da cana-de-açúcar, do fumo e da mineração, o que proporcionou o surgimento de línguas veiculares ou línguas gerais como quimbundo, “mina” e iorubá. Esse fato linguístico permitiu a convivência da

língua portuguesa com as línguas africanas, através da alternância de códigos, durante essa decorrência.

Além da questão linguística desse período no Brasil, é importante também observarmos algumas questões referentes à formação da sociedade colonial brasileira, de acordo com Pagotto (2018), que possui dois modos de interpretação da história econômica. O primeiro deles, que predominou durante o século XX, vê esse período da história brasileira como consequência de uma economia baseada no latifúndio agroexportador, tendo fora dela indivíduos sem nenhuma relevância para a sociedade, chamados de desclassificados: “O mercado interno, quando abordado, é tomado como apenas subsidiário da grande empresa agroexportadora (ou, no caso do ouro, minério-exportadora). A atividade de criação de gado de corte é um desses casos” (PAGOTTO, 2018, p. 56).

Sendo assim, essa maneira de pensar esse primeiro modo de interpretação, contribui para a hipótese da existência de uma extrema polarização linguística, em decorrência dos papéis sociais envolvendo senhores e escravos, o que teria gerado extremas consequências no decorrer dos processos de mudança, sendo que, de acordo com Pagotto (2018), no limite estaria a constituição do que é denominado de línguas crioulas.

Esse modo de pensar se diferencia da análise feita por Pagotto (2018) em relação às questões linguísticas no Brasil, durante o período colonial e o império, e a formação do que é denominado hoje de português brasileiro. Por outro lado, o segundo modelo de interpretação, que não tem sido muito explorado, sobressai dois aspectos fundamentais: (i) o papel da família como unidade econômica; (ii) a possibilidade de subsistência e geração de descendentes, que acarretam crescimento vegetativo da população (Cf. PAGOTTO, 2018, p. 57). O crescimento vegetativo, para Pagotto (2018), contribui para que a língua se perpetue através da aquisição.

Outro fator relevante é a proporção de pessoas livres no Brasil colônia, o que é destacado por Caldeira (2009), ao afirmar que o censo de 1819 indica uma população em torno de 4,39 milhões de pessoas, o que inclui nesta estimativa os índios livres registrados. Ao se retirar desse total os índios livres chega-se então a 3,59 milhões de pessoas, sendo que dessa população, de acordo com Caldeira (2009), 69,1% são livres e 31,1% de escravos. Baseado nesses dados Caldeira (2009, p. 233-234) apresenta ainda mais informações importantes, sobre os últimos

anos do século XIX no Brasil, em que havia uma média de 5 escravos por proprietários no Brasil, sendo que temos durante este período em torno de 220 mil proprietários, o que corresponde a 9% da população livre.

Outro dado bastante importante é que 2,62 milhões de livres não possuíam escravos, o que corresponde a 68% da população livre. Por esse motivo, Pagotto (2018, p. 58), questiona sobre essa população e qual seria a sua função no Brasil desse período: “Quem eram eles? O que faziam? Como ganhavam a vida? Que função desempenhavam no sistema de trocas econômicas mercantil? Que papéis tinham na sociedade?”. O período pós-colonial traz mudanças importantes no cenário histórico brasileiro.

No entanto, há poucos registros de depoimentos sobre a primeira metade do século XIX, que trazem algum esclarecimento sobre a questão da língua portuguesa falada pelos africanos e seus descendentes no Brasil. Os primeiros testemunhos sobre esse tema, que foram escritos por Saint-Hilaire (1820) e Schlichthorst (1824), segundo Silva Neto (1963 [1951]), comparam as falas dos africanos a falas de crianças, pois conservam alguma coisa de infantil, já que têm dificuldade de pronunciar o *r* e a sequência *st*. Mendonça (2012 [1933]) também cita o falar *xacôco*, referente ao português falado pelos africanos, presente em uma comédia de Gil Vicente em Portugal.

Por esse motivo, buscamos estudar uma comunidade afrodescendente remanescente de quilombo, para verificarmos comprovações de como as línguas africanas influenciaram na formação do português brasileiro. Observaremos, mais adiante, na descrição da comunidade estudada e na análise de dados da nossa pesquisa informações que ampliarão esse debate. A seguir, falaremos mais sobre essa questão do contato da língua portuguesa com as línguas africanas na formação do português do Brasil.

### 3.3 CRIOLIZAÇÃO, SEMI-CRIOLIZAÇÃO E TRANSMISSÃO LINGUÍSTICA IRREGULAR

O debate sobre a relação da língua portuguesa falada no Brasil e as línguas africanas tem atraído cada vez mais a atenção de estudiosos, principalmente quando se trata de questões relacionadas à influência que estas teriam na formação do português popular do Brasil. Inicialmente, a visão acerca dessa influência se deu

de forma pejorativa, já que ela teria “danificado”, em terras brasileiras, o português vindo da Europa, prejudicando a sua implantação, a partir da colonização do Brasil no século XVI.

Tratando-se também dessa influência, um outro debate, que se formou a partir do século XX, diz respeito à existência de um processo de “crioulização”, “semicrioulização” e, mais recentemente, de uma “transmissão linguística irregular”. Desde os anos de 1930, a discussão acerca da relação entre as línguas africanas e o português brasileiro tem se evidenciado, conforme já vimos anteriormente, principalmente em se tratando da influência daquelas na formação da variedade brasileira do português. Em relação à presença das línguas africanas durante a colonização do Brasil, é possível distinguir dois momentos, correspondentes a dois tipos de análise, um que afirma a influência africana no PB e o outro que discute a hipótese da crioulização do português brasileiros, através do contato com as línguas africanas.

Ao estabelecer uma oposição conceitual entre crioulo e semicrioulo, Elia (1979), afirma que esse último seria um estágio inicial, preparatório para o crioulo. Sendo assim, teria havido no Brasil somente semicrioulos, que têm como principal característica somente uma simplificação da língua portuguesa. O português teria, então, sofrido apenas influência das línguas africanas, já que, segundo Elia (1979), não houve uma integração de traços das línguas daquele continente, vinda para o Brasil<sup>43</sup>.

A situação de contato em terras brasileiras não produziu fusão de cultura, que teria uma relação com as línguas crioulas, provocando somente uma assimilação de uma cultura por outra, relação mais próxima de uma situação de semicrioulização. Essa influência, relacionada a questões linguísticas e culturais, dos africanos trazidos para o Brasil, é reproposta por Pessoa de Castro (1976, 1980), já que houve uma presença muito forte de falantes nativos de línguas africanas fazendo parte da população brasileira no Brasil Colônia e durante o Império.

Esse fato em si já corrobora para a hipótese de que traços linguísticos foram, de certa forma, incorporados pela língua portuguesa falada no Brasil naqueles períodos da história linguística do país. No entanto, Pessoa de Castro (2005, p. 77)

---

<sup>43</sup> Importante salientar que não havia uma teoria de sintaxe nessa época ainda. Então esses autores se valiam de uma perspectiva de gramática tradicional apenas para falar de gramática e tinham mais foco na fonética e no léxico.

também destaca “o fato de não haver sucedido um crioulo brasileiro adquirido como segunda língua ou como língua nacional, semelhante às que emergiram em outras ex-colônias americanas”, já que não há traços que levem à hipótese do português do Brasil ser considerado um crioulo, como no caso das antigas colônias inglesas, francesas e holandesas do Caribe.

Por outro lado, trabalhos recentes de linguistas norte-americanos, como Guy (1981, 1989) e Holm (1987) fizeram vir à tona o debate sobre a hipótese da criouliização no Português do Brasil. Os trabalhos de Guy (1981, 1989), dentro do quadro teórico-metodológico variacionista, analisando as diferenças entre o português popular do Brasil e o português padrão, levaram à conclusão de que uma evolução linguística natural não parece ser a responsável pelas distinções existentes entre essas variedades.

O autor afirma que um processo de criouliização teria ocorrido no passado colonial do Brasil, o que teria deixado traços no português atual falado no país. Um dos principais argumentos de Guy (1981, 1989) diz respeito a dados morfossintáticos, principalmente àqueles que se referem à concordância de número entre sujeito e verbo e entre substantivo adjetivo e determinante. Já a hipótese de Holm (1987) é de que o português do Brasil seria um semicrioulo, baseando-se em análises de dados comparativos de diferentes crioulos de base ibérica.

Através de um estudo comparando a fala dos brasileiros de estrato social mais baixo com línguas africanas e crioulas, Holm (1987) argumenta que há um paralelismo entre o português vernacular brasileiro e línguas africanas, o que seria confirmado pelas conclusões de seus estudos morfossintáticos anteriores. Sendo assim, para ele, seria consistente afirmar que o português vernacular brasileiro seria um semicrioulo, que teria resultado de um contato prolongado entre uma língua crioula e uma não crioula.

Dados sócio-históricos dão suporte para as teses defendidas pelos dois autores, principalmente os relacionados à demografia do Brasil no período da Colônia e do Império. Dos negros trazidos à força para as Américas, 40% vieram para o Brasil. Apesar de algumas divergências em relação à quantidade de africanos escravizados que chegaram em terras brasileiras, há consenso de que foram mais de três milhões, conforme já citado anteriormente por Lucchesi (2009b). Além disso, a abolição no Brasil demorou muito para acontecer, o que fez o tráfico durar muito tempo, só vindo a ser extinto em 1888, pelo menos oficialmente, haja vista os relatos

de que ainda durante certo tempo chegou a ocorrer de forma ilegal. Ainda, segundo Matos e Silva (2014), em algumas regiões houve uma concentração bastante elevada, demograficamente falando, da população afro-brasileira, vindo a superar e muito a população europeia que vivia no Brasil.

Ao pesquisar, nos anos de 1987 e 1988, uma comunidade afro-brasileira descendente de escravos, localizada no Sul da Bahia, próxima à Helvécia, também Baxter (1992), especialista em crioulos que têm como base a língua portuguesa, obteve resultados, que também já tinham sido observados por Ferreira (1969), sobre determinadas particularidades. Esses resultados evidenciam traços que não são encontrados na maior parte dos dialetos rurais e foram aprofundados por Baxter (1992, p. 31-33), sendo eles:

- (i) o uso variável de verbo derivado da terceira pessoa do singular em contextos passados, indicando estados e ações pontuais, contínuas ou habituais;
- (ii) uso de formas derivadas da terceira pessoa do singular em contextos que na língua padrão exigem o infinitivo;
- (iii) marca variável da 1ª pessoa do singular;
- (iv) negação dupla;
- (v) concordância variável de número e de gênero no SN;
- (vi) presença variável de relativos;
- (vii) presença do artigo definido no SN com referência definida.

Essas análises levaram Baxter (1992) a concluir em seus estudos que são encontrados traços no dialeto de helvécia que sugerem um processo característico das línguas crioulas, no que se refere a um processo irregular de aquisição e de transmissão de língua. Outra observação importante diz respeito a um processo de descrioulização, de acordo com Baxter (1992), que pode ser configurado pelo sistema verbal encontrado nos dialetos rurais do português brasileiro e ser derivados de dialetos como o de Helvécia.

Em vista disso, ressalta-se a necessidade de novos estudos em comunidades rurais afrodescendentes para que se pesquise a existência dessas ocorrências em outras partes do território brasileiro. Essas observações contribuirão com o mapeamento das diferenças existentes entre o português popular rural do Brasil e o português padrão, além de demonstrar os processos de mudança que se encontram nas mais diversas regiões rurais do país. Por isso mesmo, buscamos, através desse trabalho, evidências sobre a influência das línguas africanas na formação do

português brasileiro, ao pesquisarmos o preenchimento do pronome sujeito na comunidade quilombola de Lagoinha.

De acordo com Tarallo (2018b [1993]), quando se discute a provável origem crioula do português brasileiro é obrigatório passar por duas etapas, que se referem ao resgate da sua história linguística e sua história social. No caso a última tem sido usada com bastante frequência, quando se quer evidenciar hipóteses positivas a favor dessa origem crioula, o que tem sido atestado ou alegado há algum tempo por alguns linguistas, como Coelho (1880, 1882, 1886), em seu trabalho pioneiro, e Guy (1981), mais recentemente. Coelho (1880) enfatiza a semelhança de traços entre o PB e os dialetos crioulos da África, e, dentre essas, indica a falta de concordância entre artigo e substantivo, adjetivo e substantivo, por exemplo, quando ambos deveriam ir para o plural e só um o faz. Por outro lado, Coelho (1882, 1886) emite a sua própria posição admitindo que o português brasileiro não é mais degenerado do que o português europeu e que o PB está sofrendo mudanças que o distanciam do português da antiga metrópole.

Praticamente um século depois, essa discussão ainda permanecerá viva em trabalhos como o de Guy (1981), que levantam a discussão sobre as motivações que teriam determinado as mudanças históricas sofridas pelo português popular brasileiro. O questionamento se dá entre duas hipóteses: se essas mudanças teriam surgido a partir de motivações do latim ou seriam originadas como resultado de um *pidgin* falado pelos africanos escravizados que foram trazidos para o Brasil no início do século XVII e que estaria passando agora por um processo de descrioulização.

O trabalho de Guy (1981) comparou a história social do Brasil e a configuração linguística do PB com situações sociais semelhantes que se verificaram em outros locais de ocorrências de línguas *pidgins*, buscando evidências que confirmassem as suas hipóteses, argumentando que o Brasil teria todos os elementos para ter ocorrido um processo de criouliização, já que passou pelo mesmo processo verificado na Jamaica, Haiti, e na Guiana, por exemplo, alegando que a história social desses países é praticamente igual à do Brasil, em se tratando da ocorrência da formação de um crioulo. Sendo assim, para Guy (1981), seria muito mais complexo explicar a origem do português brasileiro se não tivesse surgido de um *pidgin* anterior.

Um aspecto importante a ser mencionado a respeito da formação do português brasileiro é o problema da ausência de documentos. Fala-se hoje de um

crioulo haitiano, porque há um crioulo haitiano. No caso do português, as evidências que se têm hoje são bastante indiretas. Não há documentos coloniais escritos por pessoas que viveram naquele período. Não há gravações. Na verdade, a questão é que o contexto sócio-histórico tende mais para algum tipo de influência do contato do que para o contrário. Outro fator importante a ser levado em conta é também a grande extensão territorial brasileira. Além disso, não podemos deixar de lado a democratização da escola, a expansão dos meios de comunicação de massa, os diferentes fluxos migratórios, que podem ter operado, especialmente na entrada do século XX, para uma possível descrioulização ou diminuição dos traços mais evidentes.

A procedência das evidências linguísticas de Guy (1981) está relacionada à análise de duas variáveis fonológicas, que são o apagamento do -s final, e a desnasalização das vogais finais, além de duas variáveis morfossintáticas, sejam elas a concordância no Sintagma Nominal (NPA) e concordância Sujeito-verbo (VPA). Sobre o fato dessas evidências linguísticas serem mais sugestivas do que decisivas, Guy (1981) comenta e ainda propõe três questões a serem discutidas:

The linguistic evidence, therefore, will not be decisive, but only suggestive. As we examine the findings that have been reported above, we must bear in mind the following three question:

1. Are the linguistic findings consistent with natural change process attested in the history of Portuguese and related languages?
2. Are they consistent with the reported features of creole varieties of Portuguese and related languages?
3. Are they consistent with the natural language-learning strategies which must have affected the processes of pidginization, creolization, and decreolization? (GUY, 1981, p. 291)<sup>44</sup>.

A conclusão de Guy (1981) sobre o efeito da saliência morfológica presente nas regras morfossintáticas não tem precedentes na história do português nem de qualquer outra língua românica e, em se tratando da questão da origem crioula. Por

---

<sup>44</sup> A evidência linguística, portanto, não será decisiva, mas apenas sugestiva. Ao examinarmos as descobertas que foram relatadas acima, devemos ter em mente as três perguntas a seguir:

1. Os achados linguísticos são consistentes com o processo natural de mudança atestado na história do português e das línguas afins?
2. Eles são consistentes com as características relatadas de variedades crioulas de português e línguas relacionadas?
3. Eles são consistentes com as estratégias naturais de aprendizagem da língua que devem ter afetado os processos de pidginização, criouliização e descrioulização? (tradução nossa).

outro lado, essas hierarquias de saliências seriam consideradas totalmente naturais e até mesmo esperadas:

This pattern of agreement is completely unprecedented in the history of Portuguese, and as far as I am aware, in the history of an Romance language. The development of all types of plurals in Portuguese has been governed exclusively by regular sound changes, and the syntactic agreement rules have never been influenced in any way by the mere morphological details of the words in their scope. In fact, except for the loss of NP case agreement due to phonological elimination of distinctive case-endings, the NPA and SVA rules of standard Portuguese are essentially unchanged from those from Latin. So this finding would have to be considered very surprising from the standpoint of the natural evolution hypothesis about the popular dialect. [...] From the creole origins perspective, however, these salience hierarchies would be entirely natural, even expected (GUY, 1981, p. 295-296)<sup>45</sup>.

A partir dessa análise, o que seria então o tipo de comunidade de fala crioula que teria se formado no Brasil. E, realmente, teria se formado uma comunidade crioula no Brasil? De que tipo de crioulo? Um crioulo forte? Seria um semi-crioulo? O Brasil estaria passando por um processo de descrioulização e se direcionando à língua-alvo, o PE? São muitos questionamentos que também são feitos por autores como Guy (1981), Tarallo (2018b [1993]) e outros estudiosos quando se fala da influência das línguas africanas na formação do português popular do Brasil. O fato é que não podemos negar que esse “caldeirão” linguístico do Brasil colonial não tenha deixado marcas na formação do português no Brasil.

Caso a análise de Guy (1981) esteja correta, certamente o tipo de comunidade crioula do Brasil não seria no sentido de Bickerton, já que, segundo Tarallo (2018b [1993]), as virtuais línguas-substrato, que foram trazidas do continente africano deixaram de ser faladas no Brasil e, por outro lado, não seria uma comunidade crioula do tipo *ilha*. Entretanto, esse segundo tipo de comunidade satisfaz alguns dos critérios usados para defini-la. “Também não é o caso que brasileiros nativos usem uma gama de variedades de fala intermediárias, i.e., que a

---

<sup>45</sup> Este padrão de concordância é totalmente inédito na história do Português e, tanto quanto sei, na história de uma língua românica. O desenvolvimento de todos os tipos de plurais em português tem sido regido exclusivamente por mudanças sonoras regulares, e as regras de concordância sintática nunca foram influenciadas de forma alguma pelos meros detalhes morfológicos das palavras em seu âmbito. Na verdade, exceto pela perda de concordância de caso NP devido à eliminação fonológica de terminações finais de caso, as regras do NPA e do SVA do português normal permanecem essencialmente as mesmas do latim. Portanto, essa descoberta deveria ser considerada muito surpreendente do ponto de vista da hipótese da evolução natural sobre o dialeto popular. [...] Da perspectiva das origens crioulas, entretanto, essas hierarquias de destaque seriam inteiramente naturais, mesmo se esperadas (tradução nossa).

situação etno e sociolinguística no Brasil espelhe aquela encontrada no Caribe” (TARALLO, 2018b [1993], p. 31).

A hipótese de uma comunidade do tipo *ilha*, é o que sugere a análise de Guy (1981), sendo que igualou o PB a variedades basiletais da língua, ou seja, as forças que explicam a mudança linguística no Brasil teriam o PE como língua-alvo. Apesar dos estudos já feitos para se tentar entender a tipologia das línguas crioulas, muito da história externa e interna do português brasileiro ainda permanece desconhecido. Por isso mesmo:

[...] com base no que já se conhece sobre a história interna do PB, pretendo descrevê-lo como uma língua do tipo misto: uma língua que, absurda, inesperada e estranhamente, compartilha propriedades com línguas não relacionadas, quer crioulas ou não, e que está se distanciando do superestrato original: PE (TARALLO, 2018b [1993], p. 31-32).

Problematiza-se aqui também o questionamento feito por Tarallo (2018b [1993]) acerca da hipótese crioula ainda permanecer em discussão, em se tratando da origem do português popular do Brasil. Na verdade, a questão não foi de fato esclarecida totalmente, todas as hipóteses precisam ser mais bem examinadas, mas não se pode negar, de fato, o papel do contato.

Por isso mesmo, é de suma importância buscar as raízes históricas que nos levam a percorrer as origens do português popular do Brasil, sendo que não podemos descartar nenhuma das hipóteses, mesmo que tenhamos que observar esses registros como materiais de estudo que nos auxiliem na compreensão do nosso trabalho que tem como objeto de estudo o preenchimento do pronome sujeito no português popular afro-brasileiro.

Não é possível descartar a presença de milhões de africanos em terras brasileiras e a sua grande contribuição para a formação do nosso país, em todos os setores, o que deixou também uma forte influência linguística, principalmente quando são estudadas as comunidades remanescentes de antigos quilombos. Os estudos dessas comunidades deixam bem claro que ainda restam marcas dessas influências e por isso mesmo há necessidade de um estudo maior para que se compreenda a origem do português popular do Brasil. Essas comunidades, conforme afirma Lucchesi (2015, p.107), são "verdadeiros sítios arqueológicos da história sociolinguística do país", e é nelas que se fala a variedade que os autores designam por *português afro-brasileiro*.

Como muito bem colocado por Tarallo (2018b [1993], p. 32): “A hipótese crioula do PB não é mais crucial, i.e., a criouliização pode muito bem ser colocada entre os processos de contato linguístico que ocorreram no Brasil Colonial”. Independente da origem crioula, o mais importante, segundo Tarallo (2018b [1986], p. 32) é que “o PB se presta a análises sintáticas interessantes que vão contra a hipótese de descriouliização alimentada por Guy”. Acreditamos, assim como Tarallo (2018b [1993]), ser improvável que o PE e o PB venham a se encontrar sintaticamente em um momento posterior. A língua é dinâmica e sofre influências externas e seria até ingênuo da nossa parte querer ter um controle sobre as mudanças que se sucederão com o decorrer do tempo e do espaço.

Não se deve, no entanto, deixar à parte a história externa dos estudos sobre pidgin e crioulo. Algumas colocações como as de Coelho (1880) sobre criouliização vs. formação de dialeto que, apesar de agrupar o PB com os crioulos afro-portugueses, nesse seu primeiro artigo, e em sua obra como um todo (1980, 1882, 1886), não pretende apenas fazer uma análise sobre línguas e dialetos crioulos, mas sim sobre o estado geral das línguas românicas fora da Europa. Coelho (1880), ao falar da dimensão e da população do Brasil à época, enfatiza que há um vasto campo favorável para a alteração do PE. O português brasileiro seria, então, o resultado do contato do PE com diferentes línguas e grupos étnicos, o que teria levado a essa alteração, mas que não teria levado à criouliização, e sim a uma diferenciação dialetal.

Outro ponto destacado por Coelho (1880) é que mesmo após a queda do império português, em 1822, perdurou no Brasil a tradição literária de Portugal, o que fez com que o português permanecesse bastante próximo do Brasil, pelo menos em se tratando da língua escrita. Essa afirmação nos indica claramente a comprovação a respeito da falta de evidências no que diz respeito ao contato do português com diferentes línguas e grupos étnicos neste período, já que os registros encontrados foram escritos por quem escrevia conforme a tradição literária portuguesa. De qualquer forma, Coelho (1880) também enfatiza que, desde a chegada do português europeu no Brasil, ambos os dialetos tomaram rumos diferentes. Tanto lá quanto cá, a língua portuguesa se desenvolveu e passou por alterações, já que é um organismo que se desenvolve e se altera sem que se possa ter um devido controle sobre essas mudanças que, porventura, venham a ocorrer.

Tomando como exemplo a assimetria sujeito-objeto, Tarallo (2018b [1993]) faz o seguinte questionamento a respeito do possível processo de descrioulização do PB, citado anteriormente, tendo o PE como língua-alvo:

Tomemos agora, para um breve olhar, a assimetria sujeito-objeto nos dois dialetos: o que seria necessário para o PB descrioulizar-se na direção do PE? Nem mais nem menos do que o seguinte: o PB teria literalmente que se virar pelo avesso e de ponta cabeça. Sujeitos teriam que começar a ser nulos outra vez (isto é, o PB teria que começar a readquirir suas características *pro-drop* perdidas), enquanto objetos teriam que começar a receber pronomes clíticos outra vez (TARALLO, 2018b [1993], p. 50).

Os dados históricos indicam, de acordo com Tarallo (2018b [1993], p. 50-51). que “a reversão da posição assimétrica de sujeitos e objetos foi muito lenta; pelo menos um século foi necessário para que a mudança ocorresse (1875-1982)”. Para que alguma simetria entre as posições de sujeito e objeto tivesse realmente acontecido, caso o PB tivesse passado por um processo de crioulização, teria levado pelo menos 350 anos. “Somente após 1875 deve ter havido um período que sujeitos e objetos mostravam um equilíbrio em termos de frequência das categorias preenchidas vs. vazias” (TARALLO, 2018b [1993], p. 51). Em função de quais motivos, então, por volta de, provavelmente, 1885, questiona Tarallo (2018b [1993]), o PB teria iniciado essa mudança em direção oposta ao PE?

Para Lucchesi (2015), no entanto, a miscigenação racial foi o fator principal para que o Brasil não tenha passado por um processo que caracterizasse a crioulização do português.

Desde o início da colonização, a miscigenação entre o colonizador europeu e as mulheres índias e negras foi geral e constituiu um dos vetores mais importantes da composição étnica da sociedade brasileira; ao ponto de o contingente de mulatos, no final do século XIX, atingir quase a metade da população do Brasil (LUCCHESI, 2015, p. 105).

Sendo assim, a miscigenação contribuiu para a interação sociocultural entre dominantes e dominados e trouxe um impacto muito grande, devido à grande proporção demográfica de mestiços. Esses, por sua vez, foram se integrando na sociedade dos brancos e buscando estar mais próximos da cultura do grupo dominante, incluindo, nesse contexto, uma maior aproximação da língua-alvo, como forma de ascensão social. Isso impediu que os africanos e os seus descendentes ficassem totalmente isolados da população brasileira de um modo geral, em um

mundo cultural à parte, o que seria favorável, como consequência natural, segundo Lucchesi (2009b, p. 69) para “a formação de uma língua crioula e claramente diferenciada (resultante da reestruturação independente da língua do superstrato, com forte influência do substrato)”.

O que sucedeu com a língua portuguesa no Brasil, de acordo com o que pode ser explicado através do Lucchesi (2009b), que denomina de polarização sociolinguística:

O que ocorreu, e ainda ocorre, no Brasil é um violento processo de segregação social, com evidentes reflexos linguísticos, destacando-se aí as características estigmatizadas das falas populares da população pobre (na sua maioria composta por negros e mestiços), em cuja origem certamente se encontram as mudanças operadas no bojo do processo de transmissão linguística irregular que se inicia com a aquisição defectiva do português por parte de africanos e índios. Entretanto, não ocorreu no Brasil um processo de segregação racial, de que poderia resultar uma variedade linguística autônoma de um segmento afro-brasileiro apartado; nem mesmo se pode falar hoje de uma variedade do português brasileiro característica de um segmento étnico, um português negro, por exemplo (GUY, 1981a, p.324); o que há é um português dito culto, bastante diferenciado do português popular, tanto na frequência de uso de certas variantes linguísticas (nomeadamente as regras de concordância nominal e verbal), quanto no julgamento social do uso de tais variantes, configurando o que se tem denominado a *polarização sociolinguística do Brasil* (Lucchesi 2009b, p. 69-70).

Assim, os principais fatores que contribuíram para que não ocorresse um processo de crioulação do português no Brasil foram, conforme Lucchesi (2009b, p. 69):

- (i) a proporção entre a população de origem africana e branca, que proporcionou um maior acesso à língua-alvo do que o observado nas situações típicas de crioulação;
- (ii) a ausência de vida social e familiar entre as populações de escravos, provocada pelas condições sub-humanas de sua exploração, pela alta taxa de mortalidade e pelos sucessivos deslocamentos;
- (iii) o uso de línguas francas africanas como instrumento de interação dos escravos segregados e foragidos;
- (iv) o incentivo à proficiência em português;
- (v) a maior integração social dos escravos urbanos, domésticos e das zonas mineradoras;
- (vi) a miscigenação racial.

Os fatores supracitados, de acordo com Lucchesi (2009b), explicam por que não se configurou uma língua crioula no Brasil, de uma maneira estável e representativa, já que se constatou presença muito grande de africanos na sua história sociocultural, como ocorreu em outros países em que houve também a presença maciça de uma população africana escravizada. No entanto, é importante

ressaltar que, em alguns casos mais específicos, que fogem a situações mais representativas da história sociolinguística do Brasil, ocorreram processos diferenciados. Em certas regiões do país, por exemplo, principalmente as mais afastadas, em fazendas no interior, em que houve uma grande presença afro-brasileira e em comunidades que foram oriundas de antigos quilombos, pode ter ocorrido um “intenso processo de transmissão linguística irregular em níveis muito próximos ao da criouliização” (LUCCHESI, 2009b, p. 71).

Discutiremos, a seguir, o conceito de transmissão linguística irregular, a nativização e sobre a polarização sociolinguística do Brasil.

### 3.3.1 Transmissão linguística irregular, nativização e a polarização sociolinguística do Brasil

Em se tratando de línguas crioulas, deve-se levar em conta as colônias agroexportadoras, onde surgiram grande parte delas, como o Brasil, o Caribe, o Sul dos Estados Unidos e outras regiões do continente americano, que possuem muitas semelhanças sócio-históricas. As línguas crioulas conhecidas atualmente foram resultado, segundo Lucchesi (2009a, 2015), de um processo de escravização ocorrido entre os séculos XVI e XIX, envolvendo uma grande quantidade de africanos, em torno de dez milhões, que foram sequestrados e trazidos para o continente americano pelos colonizadores europeus.

Desses africanos escravizados, estima-se que aproximadamente 40% vieram para o Brasil, o que de certa forma pode ter produzido situações de contato, como afirma Lucchesi (2009a), bem parecidas com as ocorridas no Caribe e no Sul dos Estados Unidos, que originaram as línguas crioulas. No entanto, não há nenhum registro de uma comunidade crioulófona no Brasil, sendo que estudos verificaram somente, a utilização, em algumas comunidades rurais afro-brasileiras, de um código de base lexical africana, como uma espécie de “língua secreta”, em situações muito particulares entre pequenos grupos: “Mas os antropólogos que documentaram essas línguas secretas afirmam que, nas situações normais de interação linguística, a comunidade usa uma variedade comum de português popular rural” (LUCCHESI, 2009a, p. 28).

Devido à grande quantidade de africanos que foram trazidos para o Brasil, é impossível afirmar que a situação de contato entre o português e as línguas

africanas não tenha influenciado a formação da língua portuguesa no país. Mesmo porque os afrodescendentes expandiram-se por várias regiões do Brasil, além de estarem, de certa forma, inseridos em diversos ramos de atividade econômica em todos os segmentos sociais, apesar de concentrarem-se na base da pirâmide social, devido às adversidades que enfrentaram. A população brasileira, de acordo com Lucchesi (2009a), até meados do século XIX, era constituída de apenas um terço de portugueses e seus descendentes diretos, sendo que os africanos, que predominavam na maior parte das terras brasileiras, e os índios e seus descendentes constituíam os outros dois terços restantes.

O aprendizado da língua portuguesa por esses indivíduos se deu totalmente desprovido de instrução formal acerca da língua, em condições desfavoráveis quando trabalhavam como escravizados, ficando limitado a uma comunicação que permitisse somente o entendimento para obedecer às ordens a serem cumpridas, tendo, por exemplo, um vocabulário reduzido.

Sendo assim, os filhos desses escravizados, que nasciam em condições sub-humanas, não tinham acesso à língua de seus pais que, na maioria das vezes, falavam línguas defectivas. Essas crianças, então, tinham como modelo um português falado, altamente defectivo, a partir do qual desenvolviam a sua linguagem, o seu modo de se comunicar com aqueles com quem conviviam. Apesar de essa ser uma situação prototípica de crioulização, isso não impediu que uma parte dessas crianças, filhos de africanos nascidos no Brasil, conforme afirma Lucchesi (2009a), tenham conseguido inserir-se no meio social do colonizador e tivessem acesso aos modelos de língua dominante. Como essas crianças não dispunham de um modelo fornecido pela língua nativa de seus pais, dotada de todos os mecanismos gramaticais, a transmissão linguística de uma geração a outra apresentará déficits em relação às situações normais de aprendizagem de uma língua por uma nova geração de falantes. Esse processo foi denominado como sendo uma transmissão linguística irregular:

As crianças têm de atender aos requerimentos de marcadores de tempo, modo e aspecto, de regência e ligação, operadores pronominais, etc., inerentes ao desenvolvimento de sua língua materna, a partir de dados linguísticos primários que provêm, no caso da maioria dos adultos que as cercam, de uma segunda língua desprovida da maior parte desses elementos e mecanismos gramaticais (LUCCHESI, 2009a, p. 29).

Mesmo que esse processo não provoque uma situação de crioulização, em que ocorrerá o surgimento de uma gramática original, ele produzirá certamente alterações gramaticais nas variedades linguísticas derivadas dele. Isso ocorrerá porque é justamente no momento da transmissão geracional que as mudanças gramaticais encontram seu momento crítico. O fato é que ocorreu no Brasil um processo de transmissão linguística irregular, mesmo que não tenha sido tão intenso, o que teria influenciado o surgimento das variedades do português falado por mais de um terço da população.

De qualquer modo, é relevante mencionar o contato entre línguas na formação do português brasileiro (PB). De acordo com Lucchesi (2009a, p. 29), “a grande maioria das análises publicadas sobre a **história do português brasileiro** sequer menciona o **contato entre línguas** e quando o fazem é para negá-lo”. Ele também comenta que há uma interpretação equivocada ao se utilizar testemunhos históricos do português arcaico e dados fragmentados da pesquisa dialetal portuguesa para explicar os processos de variação e mudança que afetaram a língua portuguesa no Brasil:

A **polarização sociolinguística** que marca a formação histórica da realidade linguística brasileira, apartando a fala de uma elite que sempre teve os olhos voltados para a Europa, em busca de seus modelos culturais e linguísticos, da fala da grande maioria da população que, no cadinho de sua pluralidade étnica, cultural e linguística, forjou os elementos definidores da originalidade cultural e linguística do Brasil, que tanto assombram e encantam o mundo ocidental, desautoriza todos os estudos que apresentam *uma história única para o português brasileiro*. Assim como “**o português são dois**”, a sua história é igualmente bifurcada (LUCCHESI, 2009a, p. 30).

Sendo assim, fica evidente que a realidade de uso da língua portuguesa no Brasil difere do modelo europeu de uso do português. A adoção desse modelo pelas classes dominantes do país constituiu-se como uma maneira de construir um mecanismo de exclusão social e dominação, através do preconceito com as marcas linguísticas que identificavam a pluralidade étnica brasileira. O contato entre línguas influenciou diretamente a fala da maioria da população brasileira, conforme Lucchesi (2009a), que utiliza uma norma denominada popular, e indiretamente a fala de uma minoria, privilegiada economicamente, que faz uso de uma norma denominada culta. Os mecanismos gramaticais de concordância nominal e verbal seriam aqueles que mais deixam perceber o efeito do contato linguístico “de modo que a falta de concordância constitui a grande fronteira sociolinguística da sociedade brasileira”

(LUCCHESI, 2009a, p. 31). A falta de concordância seria, então, aquela em que predomina com maior intensidade motivações para o preconceito linguístico.

Ao comentar sobre as prováveis alterações na língua portuguesa no Brasil, Silva Neto (1963 [1951]) já havia destacado a questão da polarização sociolinguística, descrevendo sobre a diferença existente dos princípios da colonização até 1808, em que era percebida claramente “a dualidade linguística entre a nata social, viveiro de brancos e mestiços que acenderam, e a plebe, descendentes dos índios, negros e mestiços da colônia” (SILVA NETO, 1963 [1951], p. 88-90).

Há também outros relatos que descrevem a mistura de línguas que marcou a formação das variedades da língua portuguesa no Brasil. Mesmo que não se possa afirmar com total convicção que ocorreram processos de crioulização durante esse período de intenso contato de línguas no Brasil, não se pode também, por outro lado, deixar de pensar nas alterações profundas no processo de aquisição do português como segunda língua, de forma irregular, por índios, que passaram pelo processo de aculturação do português, e, posteriormente, pelo grande número de negros escravizados.

Para Lucchesi (2009a), todo esse processo de transmissão linguística irregular ocorrido no Brasil, através do contato da língua portuguesa com as línguas indígenas e africanas, pode ter atingido um nível de pidginização e crioulização, mesmo que não tenham atingido diretamente o português culto falado pela alta sociedade brasileira da época, já que esses mantinham laços culturais e linguísticos muito estreitos com Portugal. Por isso mesmo, “O conceito de português afro-brasileiro fundamenta-se, não em parâmetros étnicos, mas em parâmetros sócio-históricos” (LUCCHESI, 2009a, p. 31), já que, caso o afrodescendente tenha curso universitário e seja filho de pais que possuam também nível superior, muito provavelmente será um falante da norma culta brasileira. Por outro lado, o enorme deslocamento da população brasileira da zona rural para a zona urbana, principalmente para a periferia das grandes cidades, que recebeu esse grande contingente de migrantes, fez com que o que antes era variação diatópica se tornasse variação diastrática.

O português afro-brasileiro seria designado, então, pela variedade linguística encontrada em comunidades rurais, conforme Lucchesi (2009a), que são compostas por uma população descendente de africanos escravizados, que teriam se fixado em

localidades distantes do interior do país e que estão até os dias de hoje praticando a agricultura de subsistência. A origem de muitas dessas comunidades está relacionada aos antigos quilombos de escravos foragidos, de onde foram originadas, conservando ainda certo grau de isolamento:

Dessa forma, o português afro-brasileiro guardaria uma especificidade no universo mais amplo do *português popular rural brasileiro* (ou, mais precisamente, *norma popular rural do português brasileiro*), não apenas pelas características sócio-históricas próprias às comunidades em que ele é falado, mas, sobretudo, pelas características linguísticas que o distinguiriam das demais variedades do *português popular do Brasil* (ou melhor, da *norma popular brasileira*) (LUCCHESI, 2009a, p. 32).

Podemos observar essa realidade na comunidade quilombola de Lagoinha, situada no meio rural do município de Nova Canaã, no Sudoeste da Bahia, local onde realizamos a nossa pesquisa, que é remanescente de antigos quilombos. Conforme Anjos (2000) há aproximadamente 800 comunidades rurais afro-brasileiras isoladas no Brasil. Dessas, 250 estão situadas na Bahia, não chegando a 1% da população brasileira, dentro dos 20% daqueles que ainda vivem no campo no Brasil. Com o numeroso deslocamento da população do meio rural para a periferia das grandes cidades a variação diatópica existente antes se transformou em uma variação diastrática.

De acordo com o cenário da polarização sociolinguística do Brasil, proposta por Lucchesi (2009a), a formação da norma culta se deu nos núcleos urbanos fixados no litoral, que existiam em pequena quantidade no país. Por outro lado, o português popular brasileiro se estabeleceu nas regiões de engenho, plantações de cana, mandioca e algodão, onde começavam a ampliar as atividades de mineração e a pecuária no interior do Brasil. Para lá foram enviados os africanos para a prática do trabalho escravo, onde passaram a se concentrar, juntamente com seus descendentes e a população indígena que habitava essas regiões:

Nessa interlândia multilíngue, que concentrava a maioria da população do país formado sobretudo por africanos (e indígenas) e seus descendentes mestiços e endógamos, secundados por um reduzido núcleo de capatazes, senhores, trabalhadores livres e colonos pobres de origem europeia, variedades muito alteradas da língua portuguesa iam-se formando em função da aquisição imperfeita do português como segunda língua em situações muito adversas por parte de uma grande população de falantes adultos e da socialização e nativização desse modelo defectivo de segunda língua, através de sucessivas gerações. Assim, a **norma popular brasileira** atualmente, mesmo em sua variante urbana, exhibe ainda os reflexos dos

processos de variação e mudança induzidos pelo contato entre línguas que marcaram a sua origem histórica no interior do país (LUCCHESI, 2009a, p. 33).

Esse contato, daqueles que se deslocaram das zonas rurais do Brasil para os centros urbanos, fez com que, além de terem levado os seus padrões linguísticos e culturais, adquirissem também, para que se integrassem àquele novo ambiente em que passaram a conviver, os padrões urbanos, dando origem a indivíduos com características urbanas, resultado da mistura dos dois elementos. Portanto, para se investigar o contato linguístico no Brasil e o resultado deste no português popular do Brasil deve-se observar o universo do interior do país.

No entanto, é necessário levar em conta o processo de nivelamento linguístico que tem contribuído para uma mudança no panorama sociolinguístico, por conta de diversos fatores, como a ampliação da malha rodoviária, a massificação do sistema de educação pública e a grande influência dos meios de comunicação de massa que têm contribuído para a difusão dos padrões cultos ou semicultos, conforme afirma Lucchesi (2009a).

Outro aspecto importante também a ser observado é que, ao se falar em variedade linguística, deve-se levar em conta também à sua heterogeneidade inerente. Por isso mesmo, também no caso das comunidades rurais afro-brasileiras, não se pode esperar que apresentem um comportamento linguístico uniforme:

É natural que umas comunidades tenham passado por um processo de transmissão linguística irregular mais intenso em sua formação do que outras e/ou tenham se conservado em um grau maior ou menor de isolamento. Portanto é natural que os efeitos de contato entre línguas sejam mais notados em umas comunidades do que em outras. O que se espera é que, no geral, os efeitos sejam nelas mais notáveis do que nas demais comunidades rurais, o que justificaria o conceito de *português afro-brasileiro* aqui proposto. (LUCCHESI, 2009a, p. 34).

Em relação à polarização sociolinguística do Brasil, teríamos, então, de um lado uma norma popular e, de outro, uma norma culta, ambas separadas por um grande fosso, um grande abismo que coloca, segundo Lucchesi (2009a), de um lado, uma minoria elitizada, que tem acesso e desfruta dos bens e serviços do universo da cidadania e, do outro, uma maioria, composta pela grande parte da população brasileira que é privada de acesso aos bens de consumo, aos serviços e direitos sociais básicos. Toda essa realidade começa a se configurar com a colonização do Brasil, em 1532.

Com o fim da escravidão e do Império, inicia-se também um processo de mudança do trabalho escravo para o trabalho assalariado. No entanto, afirma Lucchesi (2009b), que nada foi feito para absorver a grande quantidade de ex-escravos que agora se encontravam livres, mas sem nenhuma política social para a sua inclusão na nova sociedade que se formava, principalmente em se tratando da nova circunstância socioeconômica do país. No entanto, essa situação foi de grande importância para a história sociolinguística do Brasil. O contexto dos ex-cativos, se resume da seguinte forma:

Alguns receberam doações de terra e se dedicaram a uma cultura de subsistência, ficando à margem do grande processo produtivo. Mais marginalizados foram aqueles que abandonaram as grandes fazendas e se deslocaram ainda mais para o interior, em busca de locais ermos, onde pudessem livremente subsistir, confinando com a situação em que já se encontravam as comunidades quilombolas, de escravos foragidos. Se se dirigiam para as cidades, os ex-escravos eram mantidos nas periferias, em situação de grande miséria e marginalidade (LUCCHESI, 2009b, p. 52).

A população de negros trazidos da África, e seus descendentes nascidos aqui no Brasil, ficaram à mercê da própria sorte, abandonados pelas políticas públicas e formaram verdadeiros nichos, conservando assim as mudanças que foram decorrentes do contato radical entre línguas. Desse modo, pode-se supor a existência do uso de variedades crioualizadas do português nas comunidades rurais mais isoladas:

Como os remanescentes do “falar crioulo” descobertos, no início da década de 1960, em Helvécia, no Extremo Sul do Estado da Bahia, por Carlota Ferreira (1984), ou mesmo línguas francas de base lexical africana, das quais podem ser um reflexo as línguas secretas, descobertas na década de 1980 nas comunidades de Cafundó, em São Paulo (FRY; VOGT; GNERRE, 1981; VOGT; FRY, 1982, 1983, 1985), e Tabatinga, em Minas Gerais (QUEIROZ, 1984) (LUCCHESI, 2009b, p. 52).

O fato é que a situação socioeconômica do Brasil se manteve praticamente inalterada desde o início do projeto colonial, no século XVI, até o século XIX, sendo uma sociedade com predominância rural, com pequenos núcleos urbanos. Enquanto uma elite pequena tinha acesso à vida urbana e ao que ela oferecia a maior parte da população, formada por índios, africanos e mestiços vivia desamparada de qualquer acesso aos meios institucionais ou à cidadania, sofrendo com a cruel exploração a

que eram submetidos. Sendo assim, para Lucchesi (2009b), toda essa situação contribuiu para a formação de uma polarização sociolinguística do Brasil:

De um lado, nos restritos círculos da elite dos pequenos centros urbanos, “os grandes” da Colônia e do Império cultivavam a língua e as boas maneiras, sob inspiração dos modelos importados d’além mar. Do outro lado, nas vastas regiões do interior do país, a língua portuguesa passava por drásticas alterações, sobretudo em função do processo de **transmissão linguística irregular**, desencadeado nas situações de contato entre línguas abrupto, massivo e radical, compreendendo a aquisição precária do português por parte dos índios e africanos, a sua socialização entres esses segmentos e a sua nativização, a partir desses modelos defectivos, entre os descendentes endógamos e mestiços desses índios aculturados e africanos escravizados (LUCCHESI, 2009b, p. 52-53).

Após a extinção do tráfico negreiro, a adoção do trabalho assalariado e a imigração de uma grande quantidade de europeus e asiáticos, a situação citada acima começa a se alterar. Também a exportação do café, o que propiciou um crescimento da economia e favoreceu a industrialização nas primeiras décadas do século XX, tendo um grande avanço com o Estado Novo, na década de 1940 em diante teve um papel relevante.

Outro fator, conforme afirma Lucchesi (2009b), que contribuiu para a alteração da situação anterior a esse período, foi expansão dos centros urbanos e o surgimento na década de 1970 de megalópoles como Rio de Janeiro e São Paulo, que passam a conduzir os destinos da economia e da cultura do país. Toda essa situação impulsionou a construção de rodovias que vieram a integrar as diversas regiões do território nacional. Além disso, houve o crescimento muito rápido dos meios de comunicação de massa no país, como o rádio e, logo a seguir, a televisão, que adentraram até as comunidades mais isoladas das mais afastadas regiões brasileiras. Esse processo provocou uma profunda alteração na fisionomia do país, mesmo que não acabasse com as suas deformações estruturais, conforme Lucchesi (2009b), tendo como consequência o nivelamento, atenuando as diferenças existentes entre a fala das elites urbanas e das populações rurais pobres. Desse modo, tanto foram eliminadas marcas que caracterizavam o processo de transmissão linguística irregular, que ocorreu nos séculos anteriores, ou o desaparecimento de alguma evidência de uma provável existência de línguas crioulas, quanto permitindo que determinadas mudanças características da fala popular adentrassem na fala das camadas médias e altas da população. “Nesse caso, o grande fenômeno da imigração ocorrido entre os finais do século XIX e as

primeiras décadas do século XX pode ter desempenhado um papel decisivo” (LUCCHESI, 2009b, p. 54-55).

A vinda de um grande contingente de imigrantes europeus e asiáticos, mais de três milhões, durante o final do século XIX e o início do século XX, para o trabalho braçal assalariado no campo, conforme afirma Lucchesi (2009b), favoreceu o contato com ex-escravos africanos e com seus descendentes endógamos e mestiços. Isso ocorreu de certa forma, porque os italianos, japoneses e alemães ingressaram na base da pirâmide social brasileira.

Desse modo, a aquisição do português por esses imigrantes, na maioria dos casos, se deu através dessa parcela da população brasileira falante do português popular, com mudanças profundas, em função do processo de transmissão linguística irregular por que havia passado. No entanto, houve uma ascensão social muito rápida desses imigrantes, o que possibilitou que fossem levados para o seio da norma culta, estruturas desse português popular, adquiridas no contato inicial com a língua portuguesa.

Provavelmente, esse seja o motivo da variação de concordância de número, verbal e nominal, originadas no processo de transmissão linguística irregular, originado no processo de transmissão linguística irregular, de acordo com Lucchesi (2009b), que teria ocorrido na base da pirâmide social do Brasil colônia e do Império e se espalhando por todas as classes sociais nos dias de hoje, conforme afirma Scherre, (1994, p. 38): “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”. No entanto, é preciso salientar a grande diferença de frequência de variação, segundo Lucchesi (2009b), entre os falantes de maior poder aquisitivo e com maior grau de escolaridade e a população mais desfavorecida, socialmente falando, sem alfabetização ou pouca escolarização.

O emprego ou não das regras de concordância nominal e verbal é um dos fatores que contribui para a afirmação da polarização sociolinguística no Brasil e que demonstra claramente essa divisão entre uma elite detentora de um poder aquisitivo e nível de escolaridade mais alto em oposição a uma população com pouco acesso à escolarização e que vive na base da pirâmide social. Mesmo com as mudanças ocorridas a partir do século XX, que chegaram até as camadas mais baixas, socialmente falando, da população brasileira e que causaram significativas

modificações no cenário cultural e linguístico no país, é possível ainda perceber esse abismo que separa os dois extremos da sociedade brasileira.

A influência da cidade sobre o campo, a expansão do sistema público de ensino, mesmo que precária, e dos meios de comunicação de massa, como a televisão e o rádio têm contribuído para que marcas mais drásticas causadas pelo contato linguístico estejam em vias de desaparecimento. Esse processo pode ser percebido durante os nossos estudos, já que na comunidade estudada, a chegada da energia elétrica possibilitou também que meios de comunicação, como os citados acima, adentrassem nas casas dos moradores, além da internet, que já se faz presente nos dias de hoje. O acesso também às escolas é muito mais facilitado, tanto com a construção de unidades escolares no entorno da comunidade, como através do transporte escolar para a sede do município. De acordo com o Lucchesi (2009b), os vestígios de um processo de transmissão linguística irregular ainda se mantêm em comunidades rurais mais isoladas, principalmente naquelas que passaram por uma situação mais radical de contato linguístico e em que ocorreu uma reestruturação gramatical mais profunda.

Alguns aspectos da variação do português rural são comuns a diversas regiões, segundo Baxter (1992), outros, são particularidades de comunidades negras isoladas, resultantes de transmissão irregular do português. No Brasil, ainda há comunidades rurais originadas de antigos agrupamentos de escravos foragidos, denominados de quilombos e, como já citado anteriormente, essas comunidades são, conforme afirma Lucchesi (2015, p. 107), “verdadeiros sítios arqueológicos da história sociolinguística do Brasil” e, por isso mesmo, o estudo da fala de seus moradores é de grande importância, quando se procura explicar as variedades populares do português brasileiro, observando-se o contato entre línguas, durante a sua formação histórica. Desse modo, é possível observar essas variedades em comunidades rurais afro-brasileiras isoladas, originadas de antigos quilombos ou de ex-escravos que se juntaram em agrupamentos devido à doação de terras no final do século XIX.

Quando se fala em comunidades isoladas, é necessário lembrar que a população mais jovem tem sofrido, cada vez mais, a influência da escolarização e dos meios de comunicação como o rádio, a TV e a internet. Portanto, o que foi dito previamente é observado principalmente nos membros mais velhos, que não passaram pelo processo de escolarização ou tiveram pouquíssimo contato com este

e com os meios de comunicação já citados. Além disso, os mais jovens também têm transitado com mais frequência com os centros urbanos ao entorno da comunidade ou as grandes cidades em busca de melhores condições de trabalho e de um nível mais alto de escolarização, o que foi facilitado também pelas modernas condições de transporte. Para Lucchesi (2009b), o processo muito acelerado de penetração de empresas capitalistas no campo, durante o século passado, contribuiu e muito para que ocorressem mudanças que motivassem a descrioulização do português no Brasil.

Inicialmente, essas comunidades quilombolas viviam na clandestinidade e ilegalmente até que fosse extinta a escravidão dos africanos no Brasil, situação que se manteve praticamente inalterada até as primeiras décadas do século XX: “Tal cenário seria altamente favorável à criouliização do português, ou mesmo à manutenção, em uso corrente, de línguas francas de base africana” (LUCCHESI, 2009b, p. 56). No entanto, ao se iniciar a ocupação de grandes áreas rurais no Brasil para mineração, extração de madeira e produção de celulose, entre outros, essas comunidades foram expulsas de suas terras, em sua maioria, ou se integraram ao processo produtivo e/ou de atividades extrativistas rendendo-se à exploração da força de trabalho. Desse modo, a cultura dessas comunidades e o seu modo de vida foram afetados profundamente, o que causou o desaparecimento, em grande parte, das particularidades crioulizantes.

O contato entre línguas, durante o período de colonização do Brasil, é um dos principais momentos da história sociolinguística da formação do português brasileiro. Por isso mesmo, o português adquirido de forma precária pelos africanos trazidos para o Brasil e pelos indígenas que se integraram à sociedade colonial e, conseqüentemente, o fato dessa língua defectiva ter se tornado a segunda língua de seus descendentes e produzido um processo de transmissão linguística irregular, é o fator mais importante para o que resultou nas atuais variedades da realidade linguística brasileira:

A integração social dos escravos ou dos ex-escravos e, sobretudo, a miscigenação são fatores que atuam em dois sentidos: favorecem a assimilação dos padrões linguísticos dominantes por parte dos dominados, ao tempo em que abrem vias de introdução na fala das camadas médias e altas de estruturas criadas por mudanças ocorridas nos extratos mais baixos (LUCCHESI, 2009b, p. 71).

Desse modo, fica evidente que o processo histórico que constituiu a realidade linguística brasileira influenciou e muito o surgimento das variedades populares ocorridas no português brasileiro e as mudanças significativas em relação ao português europeu. Apesar de não ter havido um processo de crioulização, através de “uma reestruturação original da gramática e/ou a transferência de estruturas das línguas de substrato” (LUCCHESI, 2009b, p. 71), devido aos fatores da sócio-história do Brasil, que inibiram este processo, ocorreu um processo de transmissão linguística irregular do tipo leve, que é caracterizada “pela simplificação e/ou eliminação de certas estruturas gramaticais”; ou “pelo aumento na frequência de uso das formas não marcadas, bem como a sua generalização paradigmática” (LUCCHESI, 2009b, p. 71-72).

O português popular do Brasil traz, então, através das variedades que são encontradas nele, marcas evidentes desse processo de transmissão linguística irregular, como, por exemplo, a variação no emprego das regras de concordância verbal e nominal. “Porém essa simplificação morfológica, característica das situações de contato entre línguas, não se restringe, no português do Brasil, à eliminação das regras de concordância” (LUCCHESI, 2009b, p. 72).

Sendo assim, faz-se necessário uma maior compreensão da realidade linguística brasileira e da sua sócio-história, bem como do processo de polarização que marcou profundamente a formação do português brasileiro.

A compreensão da polarização da realidade linguística brasileira, bem como de sua formação histórica, desempenha, desse modo, um papel central, pois uma devida compreensão da história sociolinguística do Brasil é fundamental, desde o recorte do objeto de observação empírica até a construção final da análise qualitativa dos processos atualmente em curso na realidade linguística brasileira. Por outro lado, os resultados de análises de aspectos estruturais de variedades contemporâneas do português brasileiro, devidamente focalizadas, podem fornecer evidências empíricas decisivas para a comprovação de hipóteses históricas (LUCCHESI, 2009b, p. 72-73).

Ao analisarmos a história sociolinguística brasileira pode-se perceber que as evidências do contato entre línguas foram se diluindo, devido a diversos fatores internos e externos. Por isso mesmo, os reflexos desses contatos são percebidos nas falas de comunidades rurais afro-brasileiras, algumas originadas de antigos quilombos. Muitas dessas comunidades foram formadas com uma grande concentração de africanos, que não tinham muito acesso à língua portuguesa no

interior do Brasil. Mesmo as que não se mantêm tão isoladas atualmente, quanto no período colonial, mas que permanecem no interior do país, no meio rural, com uma população formada por afrodescendentes e remanescentes de quilombos, podem ainda revelar marcas desse contato e, por isso mesmo, pode ser de grande relevância isolar uma variedade linguística em que essa situação de contato do passado possa estar visível.

Apresentaremos, na próxima seção, uma descrição do local onde realizamos a coleta de dados para a nossa pesquisa, em que analisamos a ocorrência do sujeito nulo no português afro-brasileiro: a comunidade quilombola de Lagoinha, situada no município de Nova Canaã, na região sudoeste do estado da Bahia.

### 3.4 CONCLUINDO A SEÇÃO

Abordamos, nesta seção, o português afro-brasileiro, que, provavelmente, teria originado o português popular do Brasil, cujas variedades podem ser encontradas atualmente, de acordo com Lucchesi (2009a, 2009b) e Baxter (1992) em comunidades rurais afro-brasileiras, originadas de antigos quilombos ou de escravos que se juntaram em agrupamentos devido à doação de terras no final do século XIX.

Desse modo, expusemos inicialmente uma visão sobre o contato linguístico durante o período de colonização do Brasil, principalmente no que diz respeito à influência das línguas africanas na formação do português popular do Brasil, não deixando de mencionar também a situação inicial de contato com as diversas línguas indígenas, assim como as chamadas línguas gerais utilizadas para a comunicação com essa população que já habitava as terras brasileiras. A partir daí, são mencionados os primeiros estudos publicados sobre esse assunto, desde o final do século XIX. Para que haja um maior entendimento dessa questão, se fez necessário a abordagem do contexto histórico em que essa situação de contato linguístico ocorreu, por isso mesmo é mostrado como essa população retirada do continente africano chega a terras brasileiras e como se dá o processo de formação social do Brasil colonial.

Outro aspecto também abordado e considerado de grande importância na socio-história do português brasileiro é a formação dos quilombos, o que é citado por Lucchesi (2009b) e por Mattos e Silva (2004), durante o período colonial do Brasil.

Esses quilombos se espalharam pelo Brasil afora como forma de resistência à escravidão, durante o período escravista, sendo que ainda hoje são encontrados remanescentes desses agrupamentos, com descendentes dessas populações africanas, conforme já citado acima, e que são uma importante fonte de contribuição para o estudo da formação do português brasileiro.

Todo o contexto apresentado nesta seção nos leva a concluir que é praticamente impossível não se levar em consideração o contato linguístico entre as diversas línguas que conviveram durante o período colonial brasileiro, principalmente quando se trata do contato linguístico entre as línguas africanas faladas pelos negros escravizados, sempre em uma quantidade superior, e o português europeu, falado pelo branco vindo de Portugal, em uma quantidade demográfica inferior. Vários fatores contribuem para que essa questão seja considerada importante, desde o plurilinguismo africano até surgimento de uma língua geral africana, além do fato dos africanos serem obrigados a falar o português assim que desembarcavam em terras brasileiras, segundo Nina Rodrigues (1935).

Apesar de não se acreditar na possibilidade de ter surgido um pidgin ou mesmo ter havido um processo de criouliização, o que gera uma discussão sobre essa questão, aqui apresentada, chegando à hipótese de uma semicriouliização, transmissão linguística irregular, nativização ou, mais recentemente, de acordo com Lucchesi (2015), uma transmissão linguística irregular do tipo leve.

Todas essas discussões, de qualquer forma, são válidas e merecem ser aprofundadas, já que pouca documentação há sobre o período, sendo que através do estudo dessas comunidades rurais remanescentes de antigos quilombos ou mesmo de afrodescendentes podemos buscar pistas que nos conduzam para o que viria a ser o nosso português afro-brasileiro ou ainda como teria se formado o português popular do Brasil.

## 4 MÉTODOS

Nesta seção, abordaremos os métodos adotados para a elaboração deste trabalho de pesquisa. Iniciaremos discorrendo sobre a comunidade de fala, onde realizamos a gravação dos inquéritos, os quais são detalhados mais adiante. Logo a seguir, abordaremos o uso da língua na comunidade da Lagoinha assim como as principais dificuldades encontradas para a realização dessas gravações e a metodologia utilizada.

### 4.1 DESCRIÇÃO DA COMUNIDADE ESTUDADA

A comunidade quilombola de Lagoinha está localizada na zona rural, na parte oeste do município de Nova Canaã, a 25 km da sede municipal, a cidade de Nova Canaã, que, por sua vez, está situada a 490 km de Salvador, a 113 km de Vitória da Conquista e a 127 km de Itabuna. Nova Canaã encontra-se na Região Sudoeste do Estado da Bahia, em uma extensão territorial que faz parte do chamado Planalto de Conquista. Além disso, faz limite, ao norte, com o município de Boa Nova, ao sul, com Caatiba e Itororó, a oeste, com Poções e Planalto e, a leste, com Iguai e Ibicuí. O município possui uma população estimada em 16.482 habitantes, segundo dados do IBGE do ano de 2021, e uma área da unidade territorial de 804.618 km<sup>2</sup>.

A Lagoinha é uma comunidade já reconhecida como sendo quilombola e possui em torno de 189 moradores que vivem no local, sendo que 55 desses são membros da Associação de Agricultores Familiares da Comunidade Remanescente de Quilombo da Lagoinha. Segundo relatos dos mais antigos que ainda vivem no local, os primeiros a chegar à região vieram do sertão, das localidades de Cinzento, Lagoa D'água, Umbuzeiro, que ficam no município vizinho de Planalto, para a região das matas, em busca de terras mais férteis, com água, sendo que Manoel Nascimento Fonseca e Manoel Francisco Fonseca foram uns dos primeiros a chegar e constituir família fazendo plantações de feijão, milho, mandioca e café.

Há na região uma escola municipal de Ensino Fundamental I, até o 4<sup>o</sup> ano das séries iniciais. Por estar mais próxima de Planalto, muitos moradores se dirigem à localidade quando necessitam se deslocar para um centro urbano no seu dia a dia, para a busca de serviços de saúde, comércio, dentre outros. No entanto, o deslocamento até Nova Canaã, sede do município, também ocorre, principalmente

para frequentar a escola e estudar além das séries iniciais, ao se buscar serviços de saúde, além da realização de compras de itens básicos para uso no dia a dia.

Os adultos da comunidade geralmente não são escolarizados ou possuem pouca escolarização, sendo analfabetos ou semianalfabetos, em sua maioria. No entanto, a chegada da educação tem transformado a vida dos remanescentes de quilombolas, na comunidade de Lagoinha. Através de programas como o *Todos Pela Alfabetização* (Topa), o Governo do Estado da Bahia, em parceria com o Ministério da Educação (MEC), alfabetizou moradores adultos e idosos, de 18 a 90 anos, garantindo, assim, o acesso à educação dessas pessoas.

Na comunidade de Lagoinha, o trabalho do Topa contou com a parceria da Associação dos Agricultores Familiares Remanescentes de Quilombolas de Lagoinha, que representa os moradores e tem uma sede na região. A adesão ao programa foi a partir de 2015, numa região com alto índice de analfabetismo e dificuldade de acesso à cidade de Nova Canaã, que fica distante 25 km, conforme citado anteriormente. Os professores (alfabetizadores) são da própria região e possuem formação mínima no Ensino Médio, reunindo-se quinzenalmente para trabalhar temas como português, matemática e integração social.

Importante destacar também a criação, no ano de 2007, da já citada Associação de Agricultores Familiares da Comunidade Remanescente de Quilombo da Lagoinha que contribuiu para aumentar a autoestima dos moradores do local. Por meio da associação, houve acesso a outros programas como o Cultura Viva e Ação Griô, entre outros, gerando convites para apresentações dos remanescentes quilombolas em eventos fora da comunidade.

Dessa forma, a Associação contribuiu para a preservação das manifestações culturais da cultura da comunidade da Lagoinha, como samba de roda, samba de bumba, bumba meu boi, terno de reis, pila do café, produção de artesanato, a tradição da oralidade, entre outros. Essas ações são importantes, pois, de acordo com depoimentos dos moradores mais idosos da comunidade, há um grande desinteresse e desmotivação dos mais jovens pelas manifestações e tradições culturais e religiosas locais. Essa parcela da população local se sente atraída pela cidade, na qual busca melhores condições de educação e trabalho assalariado, oportunidades que não encontram na vida no campo, o que permite o contato com outros valores e costumes sociais.

Devemos levar em consideração, quando falamos da comunidade quilombola de Lagoinha, que o município de Nova Canaã pertencia ao município de Poções, que anteriormente pertenceu ao município de Vitória da Conquista. Sendo assim, é importante enfatizarmos que o município de Poções começou a ser povoado em torno de 1732, ou seja, a partir da primeira metade do século XVIII.

Segundo Silva (2005), essa povoação foi resultado de incursões bandeirantes em busca de ouro e pedras preciosas pelo interior da colônia. Depois, passaram a residir na região, praticando a pecuária e a agricultura, principalmente o cultivo de algodão. Silva (2005) também afirma que historiadores indicam a presença, nas terras do município de Poções, de seus verdadeiros donos, os índios Mongoiós, uma ramificação dos Camacãs, do grupo Gê, que habitavam a região muito antes da chegada dos colonizadores.

Há ainda alguns registros que indicam a vinda de negros escravizados oriundos de Angola e Moçambique e de negros nascidos no Brasil, conforme Silva (2005), que cita uma carta com data de 12 de agosto de 1780 de Manuel da Cunha Menezes, ex-governador da Capitania da Bahia, à Coroa Portuguesa, afirmando que um João Gonçalves convivia harmoniosamente num rancho com 60 pessoas, entre agregados e escravos. Outros documentos, ainda de acordo com Silva (2005), trazem relatos de que esses escravos levados para a região de Poções teriam vindos de Angola e Moçambique. Sendo assim, pode-se perceber um quadro social que era composto por brancos europeus e brasileiros, índios, principalmente mongoiós e negros escravos, tanto vindos da África quanto nascidos no Brasil.

Esse contexto sócio-histórico que encontramos na região estudada é o mesmo encontrado em todo o período do Brasil colonial. Se havia a presença dessa população negra escravizada que se deslocou para Poções, passando a habitar aquela região, que contava também com a presença de povos indígenas que, como em outras regiões do país, foram dizimados ou expulsos de suas terras, houve, de antemão, uma situação de contato linguístico entre o português, línguas africanas, indígenas e, posteriormente, mais restrita ao português e às línguas africanas, faladas pelos europeus, brasileiros e africanos.

Como em todo contexto referente à escravidão, em que aconteceram rebeliões e fugas, resultando na formação de quilombos, além do período posterior à abolição da escravatura, essa população negra se expandiu e continuou habitando a região para onde foi levada à força, originando as famílias remanescentes de

quilombos, hoje encontrados nas mais diferentes regiões do país, como na região sudoeste da Bahia, mais precisamente no município de Nova Canaã.

De acordo com relatos, a região onde se encontra o município de Nova Canaã começou a ser povoada a partir de 1910, com a chegada da família Matos, através de Bernardino Rodrigues de Matos, vindo do município de Três Lagoas, atualmente região entre Jequié e Amargosa. Houve conflitos para as demarcações de terras com os indígenas, que não aceitavam a ocupação, sendo que essas terras já pertenciam a famílias de Poções, mas não tinham demarcações bem definidas. Ao se abrir, então, a primeira clareira, de acordo com Matos e Oliveira (1999), o povoado foi se formando. A partir de 1911, chegaram outros membros da família Matos, irmãos de Bernadino, sendo Sinfrônio, Leovegildo e Inocêncio Rodrigues de Matos. A região ainda era mata fechada, no entanto era o caminho mais curto em direção ao Sul da Bahia, principalmente para as cidades de Itabuna e Ilhéus.

Em 1925, foi fundado o povoado de Água Fria (antiga Gongogi), que cresceu e foi se tornando próspero, o que atraiu muitos moradores de outros municípios e povoados vizinhos, motivados pela lavoura de café. Posteriormente, surgiu o projeto para transferir o povoado de Água Fria para um local que fosse mais acessível, quando foi comprada a Fazenda Caldeirão, com o objetivo de se construir o povoado e transformá-la em cidade. Isso ocorreu em 1961, quando finalmente Água Fria se separa e se emancipa politicamente do município de Poções, através da Lei 1.540 de 21 de outubro de 1961, publicado no Diário Oficial 08 de novembro de 1961, e é criado, então, o município de Nova Canaã.

A seguir, expomos algumas fotografias tiradas durante o processo de coleta de dados na comunidade. Discorremos, logo após, sobre o estudo do uso da língua dentro de uma comunidade de fala.

**Figura 2** - Moradores da Comunidade da Lagoinha, após reunião da Associação.



**Figura 3** - José Carlos Assunção Novaes com Eliete Fonseca Santos e Lídio Fonseca Araújo, moradores da Comunidade da Lagoinha



**Figura 4** - Reunião da Associação



**Figura 5** - Moradora mostra forma artesanal de produção de farinha na comunidade



**Figura 6** - Sede da Associação de Moradores da Lagoinha



**Figura 7** - José Carlos Assunção Novaes e moradores da comunidade



#### 4.2 O ESTUDO DO USO DA LÍNGUA DENTRO DA COMUNIDADE DE FALA

Toda pesquisa que tenha como intuito a análise dos problemas da linguagem humana irá se deparar com o desafio, de acordo com Guy (2007 [1993]), de olhar ao mesmo tempo, para a organização das formas linguísticas de um lado e para a sua significância social de outro. Apesar de seus encantos, essa dualidade, em se tratando da pesquisa linguística, torna-se também a origem de determinados problemas metodológicos que são cruciais dentro dessa área.

As pesquisas sociolinguísticas têm mostrado que não há necessidade de amostras tão grandes como as usadas em outras pesquisas de natureza social (de intenções de voto, por exemplo) para se analisar fenômenos variáveis, uma vez que o uso linguístico é mais homogêneo do que o comportamento humano acerca de outros fatos, em virtude de não estar tão sujeito à manipulação consciente (com a ressalva de que no caso dos estereótipos possa haver algum grau de manipulação consciente). Assim, por exemplo, numa dada comunidade, resultados obtidos a partir da análise da fala de cerca de 60 informantes serão apenas ratificados em amostras maiores, ou seja, a partir de certo número de informantes os resultados tornam-se redundantes (COELHO *et al*, 2015, p. 114).

A metodologia quantitativa vem se amparando, cada vez mais, do uso de tabelas e gráficos, usados para a apresentação de dados, além de indicadores estatísticos para resumo e inferência, além de testes de significância e confiabilidade e técnicas analíticas quantitativas. “Esses desenvolvimentos têm se pronunciado

particularmente na área dos estudos de comunidade de fala e variação sociolinguística” (GUY, 2007 [1993], p. 20).

Guy (2007 [1993]) resume as três fases de uma análise quantitativa da seguinte forma: (i) coleta de dados; (ii) redução e apresentação de dados; (iii) interpretação e apresentação de dados. Ao abordar a questão referente à pesquisa sociolinguística, Silva-Corvalán (1988) também cita alguns possíveis passos a serem seguidos, afirmando que, ao mesmo tempo, não há uma receita exclusiva para se fazer uma investigação linguística. Sendo assim, a sequência abaixo é apresentada por Silva-Corvalán (1988, p. 16):

1. Observação da comunidade e hipóteses de trabalho.
2. Seleção dos falantes.
3. Coleta dos dados.
4. Análise dos dados:
  - 4.1. Identificação da variável.
  - 4.2. Identificação dos contextos.
  - 4.3. Codificação.
  - 4.4. Quantificação e aplicação de procedimentos estatísticos.
5. Interpretação dos resultados<sup>46</sup>.

O que podemos observar, na sequência proposta por Silva-Corvalán (1988), é a apresentação de dois itens anteriores ao item 3 (Coleta de dados) e que não aparecem no resumo das três fases da pesquisa quantitativa apresentada por Guy (2007 [1993]). Os itens 1 e 2, referentes à observação da comunidade e hipótese de trabalho e seleção dos falantes, respectivamente, são de grande relevância. É importante enfatizar esse aspecto em função dos rumos que nortearam esta pesquisa linguística.

Também é importante salientar que a ordem dos passos acima é flexível, conforme afirma Silva-Corvalán (1988), já que é possível, por exemplo, completar a seleção de falantes depois da coleta dos dados, assim como há ainda a possibilidade de que a hipótese inicial seja reformulada durante o processo de análise dos dados, podendo levar a uma nova seleção de falantes ou a uma outra etapa de coleta de dados. Silva-Corvalán (1988, p.1 7) também afirma que:

---

<sup>46</sup> Tradução nossa.

Además, aunque frecuentemente se inicia una investigación con una hipótesis específica, es posible también aproximarse al estudio de una comunidad de hablantes con una hipótesis muy general, como por ejemplo, que dada la evidencia aportada por otros estudios, el habla de los individuos de dicha comunidad se caracterizará por la existencia de ciertos rasgos lingüísticos variables cuya frecuencia de ocurrencia estará condicionada por factores independientes de tipo lingüístico y/o extralingüístico.<sup>47</sup>

De qualquer forma, um dos problemas mais importantes a serem resolvidos é qual a comunidade a ser estudada e quais os falantes serão escolhidos, conforme Silva-Corvalán (1988), de forma que a amostra seja representativa dessa comunidade escolhida. A partir do momento que a comunidade é definida, a observação e a participação em interações diádicas e em grupo permitirão ao sociolinguista identificar uma ou mais variáveis.

Quanto ao segundo passo, que diz respeito à seleção da amostra de falantes, isso vai depender, conforme Silva-Corvalán (1988), do objetivo do estudo e da hipótese inicial:

Por ejemplo, los factores variables independientes que proponamos como condicionantes de la variación, tales como sexo, edad, nivel educativo, etc., deberían estar representados en la muestra. Si asignamos más importancia a los factores lingüísticos y los consideramos el foco central del estudio, podemos incluir una muestra homogénea de hablantes en cuanto a edad y nivel educativo, por ejemplo. En cualquier caso, la población del estudio debe seleccionarse siguiendo un método que asegure una muestra representativa, la que se puede lograr haciendo uso de las técnicas de muestreo desarrolladas para las ciencias sociales (SILVA-CORVALÁN, 1988, p. 17)<sup>48</sup>.

Após o estabelecimento dos critérios referentes aos fatores sociolinguísticos citados previamente, como sexo, idade, nível de escolaridade, classe social, local de

---

<sup>47</sup> Além disso, embora a pesquisa seja frequentemente iniciada com uma hipótese específica, também é possível abordar o estudo de uma comunidade de falantes com uma hipótese muito geral, como por exemplo, que dadas as evidências fornecidas por outros estudos, a fala dos indivíduos da referida comunidade será caracterizada pela existência de certos traços linguísticos variáveis cuja frequência de ocorrência será condicionada por fatores independentes do tipo linguístico e/ou extralinguístico (tradução nossa).

<sup>48</sup> Por exemplo, os fatores variáveis independentes que propomos como fatores condicionantes de variação, como sexo, idade, escolaridade etc., devem ser representados na amostra. Se dermos mais importância aos fatores linguísticos e os considerarmos o foco central do estudo, podemos incluir uma amostra homogênea de falantes em termos de idade e escolaridade, por exemplo. De qualquer forma, a população do estudo deve ser selecionada seguindo um método que garanta uma amostra representativa, que pode ser alcançada utilizando as técnicas de amostragem desenvolvidas para as ciências sociais.

origem, entre outros, segundo Silva-Corvalán (1988), é possível usar a técnica de amostra aleatória para obter uma amostra representativa de cada grupo.

Depois de escolhida qual comunidade será estudada, deve-se partir para o próximo passo que é a coleta de dados linguísticos. Durante essa fase, os questionamentos propostos por Guy (2007 [1993]) e suas respectivas respostas devem ser levados em conta para que não tenhamos dados meramente obtidos, sem nenhuma preocupação com a pesquisa linguística em si. Assim, é necessário ter cuidado com a busca de um resultado o mais próximo possível da realidade linguística local. Para isso, é importante observar respostas para as seguintes perguntas:

Como obtemos os dados?  
 Os dados são válidos para refletir o fenômeno investigado?  
 Os dados da amostra são representativos da população em geral?  
 Os procedimentos para a obtenção dos dados são confiáveis e reproduzíveis?  
 O que pode ser feito para minimizar a parcialidade dos dados? (GUY, 2007 [1993], p. 20)

Depois de feitas essas reflexões e observados os devidos procedimentos para o início da coleta de dados, de acordo com uma metodologia possibilite um bom aproveitamento do material pesquisado, segue-se para a gravação das entrevistas. Um fator preponderante para se iniciar a pesquisa sociolinguística e obter êxito nas gravações é o contato com os falantes e a possibilidade de conseguir que eles se disponham a conversar livremente durante uma sessão de gravação.

Por isso mesmo, há, segundo Silva-Corvalán (1988), uma grande dificuldade em encontrar os indivíduos certos para a pesquisa e como dar uma explicação adequada a cada um deles para que aceitem participar de uma conversação gravada. Sendo assim, o mais recomendável é que se faça um contato prévio para que se evite perder tempo com pessoas que não tenham interesse em participar do estudo linguístico.

Isso só é possível, de acordo com Silva-Corvalán (1988), adquirindo a confiança da população a ser estudada para que o investigador possa realizar uma conversação gravada de mais ou menos uma hora. Essa técnica, denominada entrevista sociolinguística, após o pioneirismo de Labov (1966), de acordo com Silva-Corvalán (1988), tem sido modificada e aperfeiçoada desde então. O principal

objetivo da conversação gravada é obter uma amostra de fala casual, natural, que se aproxime o máximo possível de uma fala vernácula espontânea, semelhante à utilizada pelo entrevistado em sua vida diária.

A principal motivação para esse objetivo, conforme Silva-Corvalán (1988), é que quanto mais o falante se concentra no conteúdo do que está falando e deixa de prestar atenção à sua fala, os dados referentes ao estilo casual e vernáculo são mais sistemáticos e regulares. Esta situação leva a uma questão, referida por Labov (2008 [1972]) como o paradoxo do observador. Como pode o observador conseguir uma fala casual e vernácula, que naturalmente se usa na intimidade, entre amigos e familiares, quando se está em uma situação artificial e com estranhos? De acordo com Labov (2008 [1972], p. 244), “O objetivo da pesquisa linguística na comunidade deve ser descobrir como as pessoas falam quando não estão sistematicamente observadas – no entanto, só podemos obter tais dados por meio da observação sistemática”.

Esse não é um problema sem solução, de acordo com Labov (2008 [1972]), bastando que o entrevistador se utilize de diversas técnicas que podem ser incorporadas à entrevista, como tentar desviar a atenção do falante, tentar deixar o entrevistado à vontade de forma que não se sinta em um ambiente de entrevista, envolver a pessoa em perguntas que envolvam emoções fortes, lembranças do passado, entre outras.

Durante a nossa pesquisa, procuramos estar atentos a todos os problemas encontrados, quando se trata de um estudo linguístico. Por isso, a seguir, discorreremos sobre a metodologia empregada neste estudo e como se deu esse processo.

## 4.3 METODOLOGIA

### 4.3.1 Amostras de falas: entrevistas

O *corpus* desse estudo provém de gravações feitas entre 2018 e 2019, com seis informantes da comunidade quilombola de Lagoinha, situada no município de Nova Canaã, no Sudoeste da Bahia, todos da zona rural, analfabetos ou semianalfabetos, distribuídos em três faixas etárias: de 25 a 35 anos, de 45 a 55 anos e de 65 anos em diante. O espaço de aproximadamente uma geração entre as

faixas etárias permite uma observação em *tempo aparente*, o que nos possibilita fazer um diagnóstico sobre o processo de variação e mudança. Cada faixa etária é composta por um informante do sexo masculino e um do feminino. Acreditamos que o quadro que se formará a partir da fala de informantes com baixo nível de escolaridade, usuários do português popular do Brasil em seu dia a dia, no trabalho, no contato com a família e com os amigos, será bastante revelador do estágio em que se encontra o processo de mudança que está sendo analisado em nossos estudos. Como sabemos, o uso do sujeito pronominal ou do sujeito nulo não está associado a nenhum estigma, por isso esperamos que os resultados sejam um reflexo real da situação da variável.

Sendo assim, este trabalho se caracterizou pelo seu caráter bibliográfico e de pesquisa de campo e compreendeu as seguintes etapas:

- a) estudos sobre aspectos teórico-metodológicos da Sociolinguística, da Teoria Gerativa e da história do PB;
- b) levantamento e seleção de pesquisas linguísticas disponíveis sobre o tema;
- c) pesquisa de campo, com gravação de inquéritos, obedecendo à orientação da metodologia sociolinguística, com seis informantes, assim distribuídos em três células de dois informantes cada:

(i) Faixa Etária:

Faixa 1	25 a 35 anos
Faixa 2	45 a 55 anos
Faixa 3	acima de 65 anos

(ii) Sexo:

Masculino;

Feminino.

- d) análise do material de pesquisa: para a análise do material, utilizamos os recursos do Programa Goldvarb X, de forma a considerar frequência e peso relativo dos fatores das variáveis que foram estudadas;

e) considerações sobre a observação feita e a sócio-história da região.

Buscando verificar a implementação do uso do sujeito pronominal, fazemos uma comparação entre os resultados encontrados e os trabalhos de Duarte (1995, 2019), sobre o parâmetro do sujeito nulo na norma culta do português brasileiro; Lucchesi (2004, 2009d); sobre o português afro-brasileiro; e nossa pesquisa anterior, realizada em 2007 e publicada em 2021, sobre português popular do interior do estado da Bahia, respectivamente.

Acreditamos que o quadro que se formará a partir da fala de informantes com baixo nível de escolaridade, usuários do português popular do Brasil em seu dia a dia, no trabalho, no contato com a família e com os amigos, será bastante revelador do estágio em que se encontra o processo de mudança que está sendo analisado em nossos estudos.

#### 4.3.2 Caracterização dos informantes

Partindo do princípio de que os falantes do português popular do Brasil caracterizam-se por possuir um baixo nível de escolaridade, procuramos na escolha dos informantes selecionar aqueles que eram analfabetos ou semianalfabetos, possuindo uma formação primária incompleta, buscando analisar a utilização do sujeito pronominal nos falantes do português popular afro-brasileiro do Brasil.

Com base em uma ficha do informante aplicada no trabalho de campo, foi montado o quadro seguinte:

**Quadro 1.** Informantes da comunidade quilombola de Lagoinha (município de Nova Canaã – BA).

	<b>CÓDIGO</b>	<b>SEXO</b>	<b>IDADE</b>	<b>ESCOLARIDADE</b>	<b>OCUPAÇÃO</b>	<b>INFORMANTE</b>
FE	SL01	M	30	semianalfabeto	lavrador	E.F.S.
I	SL02	F	32	semianalfabeto	lavradora	M.S.D.A.L.
FE	SL03	M	45	semianalfabeto	pedreiro	A.F.G.
II	SL04	F	49	semianalfabeto	lavradora	V.F.J.R.
FE	SL05	M	90	analfabeto	lavrador	L.F.A.

III	SL06	F	81	analfabeta	lavradora	A.R.J.
-----	------	---	----	------------	-----------	--------

Fonte: Elaborado pelo autor

Podemos observar no Quadro 1 que os informantes da comunidade quilombola de Lagoinha exercem atividades rurais. Todos mantêm uma relação com a zona rural e convivem nesse ambiente.

#### 4.3.3 Constituição do *corpus*

Todo trabalho de pesquisa que tem como base a variação linguística requer uma amostra de dados para ser analisada. Uma vez definido o fenômeno a ser estudado, deve-se buscar a seleção dos informantes que serão observados e, através da definição da variável dependente, levantar hipóteses, para analisá-la através das variáveis independentes ou grupo de fatores. É importante que essa seleção atenda às hipóteses levantadas no início da pesquisa e represente as variáveis sociais observadas.

Nesta pesquisa, foram utilizados como amostra inquéritos gravados na comunidade quilombola de Lagoinha, no município de Nova Canaã, na região Sudoeste da Bahia. Para isso, foi necessário um primeiro contato com a comunidade, através da Secretaria de Educação do município, que nos indicou Romilce Rodrigues, coordenadora do Programa Todos pela Educação – Topa, que já mantinha contato com a Associação de Remanescentes de Quilombo da Lagoinha, através da coordenação do programa e como incentivadora do resgate de manifestações culturais na localidade.

Essa ação foi necessária, devido à necessidade de uma intermediação de alguém que tivesse acesso à comunidade para facilitar o contato com os moradores do local, em um primeiro momento. Outro fator também seria em relação à localização geográfica da comunidade, em uma região de difícil acesso, o que seria complicado sem alguém que conhecesse as vias de acesso à localidade. A primeira visita à comunidade se deu durante uma reunião da Associação de Agricultores Familiares da Comunidade Remanescente de Quilombo da Lagoinha, com o acompanhamento de um professor do município de Nova Canaã, José Nilton, que atuava na região, e do professor de História de escolas do Ensino Médio dos

municípios vizinhos de Iguai e Ibicui, Rodrigo Almeida, que também pretendia desenvolver também uma pesquisa na comunidade, o que não foi levado adiante.

Essa foi uma visita somente para conhecer a região e a comunidade, mas os moradores questionaram sobre o motivo da visita e falaram sobre o interesse de algumas pessoas com a comunidade, mas que depois não retornavam e não davam continuidade ao trabalho.

Sendo assim, foi necessário explicar à comunidade que havia interesse em realizar uma pesquisa e que, caso isso ocorresse, haveria outros contatos e esclarecimentos com a associação de moradores da Lagoinha. A princípio, não houve muito interesse por parte da comunidade em participar da pesquisa e precisamos retornar outras vezes, sempre com o intermédio de Romilce Rodrigues, que nos auxiliou passando algumas informações e indicando as pessoas que deveríamos conversar para que fossem quebradas as barreiras que impediam uma maior aproximação com a comunidade.

Depois de algumas reuniões, principalmente com a presidente da associação, na época, Eliete Fonseca Santos, e os membros mais idosos, ficou estabelecido que poderiam ser feitas as gravações na localidade e então partimos para a busca dos informantes, de acordo com as características estabelecidas, que eram três faixas etárias (25 a 35 anos; 45 a 55 anos; acima de 65 anos) e dois sexos (masculino e feminino).

#### 4.3.4 Gravação dos inquiridos de fala

Depois que as características dos informantes foram determinadas, os primeiros contatos foram estabelecidos com os informantes. Havia uma grande dificuldade em realizar as gravações, pois o local era de difícil acesso, além da pouca disponibilidade dos informantes para que fossem feitas as entrevistas. Muitos também se mostravam indiferentes ou indispostos a participar, mesmo que anteriormente tivessem aceitado participar da pesquisa, através da conversa gravada.

Dessa forma, todas as orientações foram passadas para a presidente da Associação de Agricultores Familiares da Comunidade Remanescente de Quilombo da Lagoinha, Eliete Fonseca dos Santos, para que entrasse em contato com os informantes e que também ficasse responsável para realizar as gravações das

conversas, sendo previamente orientada e auxiliada pelo autor desta pesquisa. Esse procedimento facilitou o desenvolvimento da gravação dos inquéritos e resolveu o problema do *paradoxo do observador*, citado anteriormente, segundo Labov (2008 [1972]), já que os moradores ficaram mais à vontade quando estavam sendo entrevistados por alguém da sua própria comunidade.

Depois que foram definidas as características dos informantes e feitos os primeiros contatos para que fossem marcados os locais e horários, as gravações foram acontecendo nas residências, nos momentos de descanso, em que se encontravam com os familiares, nas salas, varandas, quintais, ou mesmo nos locais de trabalho na “roça”, como por exemplo, durante a colheita de café. Em um primeiro momento, foram passadas as informações para a utilização do gravador e as orientações básicas de como a inquiridora deveria proceder, sem interferir muito para que as entrevistas fossem feitas o mais naturalmente possível. A seguir, foram iniciadas as gravações, através da interação da documentadora com os informantes, sempre orientada e auxiliada pelo autor desta pesquisa.

Mesmo que os informantes soubessem que estavam sendo gravados, houve uma intenção, por parte do inquiridor, de tornar o ambiente o mais próximo da realidade do dia a dia, buscando-se uma fala mais casual. As conversas versaram sobre a vida na comunidade, as festas populares e religiosas, o passado, a infância, brincadeiras, lendas, cantigas folclóricas, narrativas de experiências pessoais, algum acontecimento trágico que tenha ocorrido na comunidade, na vida do informante ou com alguém de sua família, entre outras que possibilitavam deixar o informante à vontade para falar sem ser interrompido, deixando-o mais próximo de sua vivência e de seu cotidiano, conforme a técnica utilizada para gravação de inquéritos na metodologia sociolinguística e temas apresentados por Lucchesi (2009c, p. 161), adotando a técnica sugerida por Labov (2008 [1972]).

Dentro desse contexto, as narrativas de experiências pessoais são consideradas as que mais tendem a evoluir para uma conversação espontânea, segundo Labov (2008 [1972]). Desse modo, o falante se envolve mais com o conteúdo da conversa deixando de se preocupar com a fala em si. Entre esses, está o assunto relacionado a risco de vida, que tem se revelado um dos principais, de acordo com como a resposta a perguntas do tipo: “Você já viveu uma situação em que correu sério risco de morrer?” (LABOV, 2008 [1972], p. 245).

No caso da escolha dos temas, foram de acordo com a realidade da comunidade e de acordo com o direcionamento da conversação, no sentido de deixar o informante bem à vontade, conforme citado acima. Caso houvesse necessidade, esses temas eram inseridos na conversa para que o relato fosse prosseguido e a gravação não fosse interrompida. Foram realizadas entrevistas da maneira mais informal possível, deixando que os informantes guiassem o curso da conversa, de acordo com o seu interesse. Ao abordar esses temas, buscou-se também informações que pudessem ser úteis na análise das variáveis extralinguísticas, que são muito importantes no estudo da língua de uma determinada comunidade. As gravações duraram entre 45 e 60 minutos, a depender do informante.

#### 4.3.5 Processamento quantitativo dos dados

Depois de executadas as seis entrevistas com falantes do português afro-brasileiro na comunidade de Lagoinha, no município de Nova Canaã, Bahia, passou-se à audição e transcrição dos áudios. A transcrição foi feita pelo autor da pesquisa. Foi utilizada a transcrição grafemática, já que foram considerados como relevantes os fatos linguísticos relacionados à fala dos informantes.

Para que a análise dos dados seja feita, é necessário definirmos a variável em estudo, ou seja, a variável dependente, e determinarmos os contextos de ocorrência, através das variáveis independentes. A partir daí partimos para a codificação dos dados, uma vez determinada a variável dependente, a realização do sujeito pronominal, e definidas as variáveis independentes/explanatórias ou grupos de fatores.

Logo após foram levantadas as ocorrências, no *corpus* constituído para a análise quantitativa da presença ou ausência do sujeito pronominal nas orações constituídas por verbos finitos. A partir daí, essas ocorrências foram codificadas e processadas no programa computacional de regras variáveis Goldvarb X, que aponta frequência e pesos relativos, de extrema importância para a interpretação dos dados e é utilizado especificamente para o estudo da variação linguística.

Desse modo, seria possível a verificação dos resultados de que disporíamos para os estudos, apresentados através de tabelas e gráficos na nossa análise dos dados. Além da porcentagem de ocorrências, também se estabeleceu um peso

relativo para definir a interferência de cada fator linguístico ou extralinguístico, considerando-se a ação simultânea de todos os fatores que afetam a produção verbal do falante no momento de cada ocorrência. Sendo assim, partimos para a análise dos dados obtidos, apresentada a seguir.

#### 4.4 CONCLUINDO A SEÇÃO

Nesta seção, apresentamos a comunidade quilombola de Lagoinha, situada no município de Nova Canaã, Bahia, local onde realizamos a gravação de nossos inquéritos, que formaram o *corpus* da pesquisa realizada neste trabalho. Sendo assim, falamos também sobre as origens da comunidade e da origem e sócio-história dos municípios de Nova Canaã e de Poções.

Logo a seguir, abordamos a importância da pesquisa quantitativa para a pesquisa linguística. Nesta seção, também mostramos a metodologia empregada para a coleta de dados até a interpretação desse material coletado e analisado posteriormente. Por fim, apresentamos a metodologia utilizada, mostrando as amostras de fala, como foram realizadas as entrevistas, as informações acerca dos informantes, como foi constituído o *corpus* da pesquisa, o modo como foram realizadas as gravações dos inquéritos e o processamento quantitativo dos dados.

## 5 ANÁLISE DE DADOS

Neste capítulo, apresentamos os resultados relativos ao fenômeno de preenchimento do sujeito pronominal. Para isso, exibimos os fatores linguísticos e extralinguísticos que condicionam a presença do sujeito pronominal no português afro-brasileiro na comunidade quilombola de Lagoinha, localizada na região sudoeste do estado da Bahia. Assim, apresentamos os contextos que constituem o encaixamento estrutural e social da realização do sujeito pronominal na comunidade de fala estudada.

Para realizar a análise desse fenômeno variável, foram observados os seguintes grupos de fatores linguísticos:

- a) pessoa do discurso;
- b) desinência gramatical do verbo;
- c) caracterização semântica do sujeito;
- d) tipo de oração;
- e) tipos de verbo;
- f) menção anterior do sujeito;
- g) presença de constituinte à esquerda do verbo.

Em relação aos fatores extralinguísticos, consideramos as seguintes variáveis sociais:

- a) sexo;
- b) faixa etária;
- c) grau de escolaridade;
- d) estada fora da comunidade.

Na realização desta análise, submetemos os dados extraídos de inquéritos de fala de seis informantes ao programa Goldvarb X<sup>49</sup>, o qual atribui peso relativo a cada fator das variáveis apontadas como estatisticamente relevantes, mensurando a interferência simultânea de todos os fatores, com a finalidade de verificar correlações entre os fatores e a realização do fenômeno variável em análise, que, nesse caso, é a realização do sujeito pronominal. Todas as variáveis linguísticas e

---

<sup>49</sup> Disponível em: <http://individual.utoronto.ca/tagliamonte/goldvarb.html> Acesso em: 10 mar. 2022.

extralinguísticas, apresentadas anteriormente, foram consideradas nesta investigação<sup>50</sup>.

Após a submissão dos dados à análise, não foram considerados significativos, em termos estatísticos, os seguintes grupos de fatores estruturais: (i) tipo de orações; (ii) tipos de verbo; (iii) presença de constituinte à esquerda do verbo. Já em relação às variáveis sociais, não se mostraram estatisticamente significativas as seguintes variáveis: (i) faixa etária, (ii) grau de escolaridade e (iii) estada fora da comunidade.

Nesse caso, somente uma variável social foi considerada estatisticamente significativa: a variável sexo. No que tange especificamente à escolaridade, acreditamos que os resultados possam ser explicados em função de os entrevistados de nossa amostra serem analfabetos ou semianalfabetos, isto é, com nenhum ou baixo grau de escolaridade. O mesmo ocorre em relação ao grupo de fatores estada fora da comunidade em que a diferença entre os fatores controlados já parecia não ser, desde o princípio, muito relevante: o tempo passado fora da comunidade era curto, sendo para todos os informantes maior o tempo de convivência na zona rural.

Os resultados da análise multivariada, na qual consideramos todas as variáveis linguísticas e extralinguísticas encontram-se listados a seguir. Vale ressaltar que, para realizarmos essa análise, todos os *knockouts*<sup>51</sup> verificados na análise anterior (análise com valores percentuais) foram excluídos desta rodada. As variáveis apontadas como estatisticamente significativas foram: 1 menção anterior do sujeito; 2 sexo; 3 pessoa do sujeito; 4 caracterização semântica do sujeito; 5 desinência gramatical do verbo.

A partir dos resultados gerais expostos, apresentamos, em um primeiro momento, na subseção 5.2, a análise multivariada, que considera o peso relativo das variáveis linguísticas sobre o fenômeno de preenchimento do sujeito pronominal, na mesma ordem indicada anteriormente (ordem de seleção das variáveis) com

---

<sup>50</sup> Neste trabalho não foram contabilizados os sujeitos em orações coordenadas que eram os mesmos da oração anterior, uma vez que, nesses casos, geralmente não há necessidade ou obrigatoriedade de realização do segundo sujeito em línguas de sujeito obrigatório.

<sup>51</sup> Ocorrem casos de *knockout* quando fatores apresentam aplicação categórica (100% ou 0%). Dessa forma, como a análise multivariada só leva em conta casos de variação, é necessária a exclusão desses dados.

exceção da variável *sexo*, analisada à parte, na subseção 5.3, por se tratar de uma variável extralinguística. A seguir, analisaremos os dados, detalhando cada variável e apresentando exemplos ilustrativos retirados do *corpus* de nossa pesquisa, iniciando com a análise da variável dependente.

## 5.1 A VARIÁVEL DEPENDENTE

Neste estudo, a variável dependente considerada foi a realização do sujeito pronominal. Foram depreendidas, no total, 1200 ocorrências de sujeitos pronominais realizados ou não, como exemplificado abaixo:

- (1) *Nós* fazia bola de saco. (SL01)
- (2) *Eu* morei uns tempo em São Paulo (SL03).
- (3) Ø Somos três. (SL02)
- (4) Ø Não tô vendo nada disso. (SL05)

Em (1) e (2) temos exemplos de casos em que o sujeito pronominal se encontra expresso, casos da variante realização, portanto. Já em (3) e (4), por sua vez, temos exemplos de casos de sujeito nulo.

Após a realização da análise multivariada, levando em conta todas as variáveis linguísticas e todas as variáveis sociais, em que todos os *knockouts* encontrados na análise anterior (com percentuais) foram excluídos, sendo que foram extraídos apenas cinco dados relativos aos *knockouts*, apresentamos os resultados a partir de agora. Na Tabela 1, a seguir, exibimos a frequência geral da ocorrência das variantes na amostra investigada.

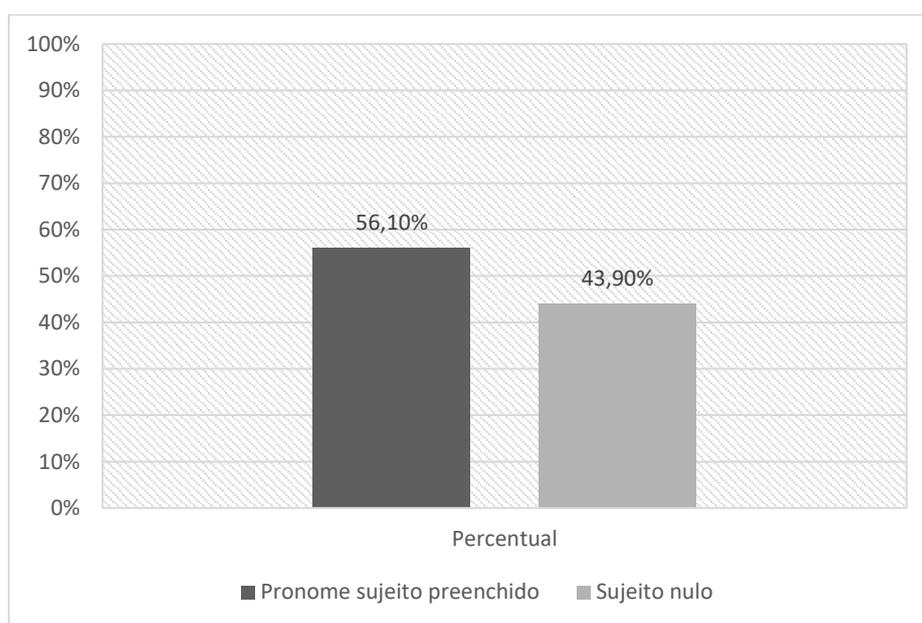
**Tabela 1** - A realização do sujeito pronominal no português afro-brasileiro na comunidade quilombola de Lagoinha – Nova Canaã – BA

Variantes	Aplicação/Total	Percentual
Pronome sujeito realizado	673/1199	56,1%
Sujeito nulo	526/1199	43,9%

Fonte: Elaborado pelo autor

Conforme os resultados exibidos na Tabela 1, a frequência geral de realização fonética do pronome sujeito, no total dos dados em análise, foi de 56,1% (673/1199) de sujeitos pronominais expressos em oposição a 43,9% (527/1199) de sujeitos nulos. Podemos visualizar melhor esses resultados, observando o gráfico apresentado na Figura 8.

**Figura 8** - Índices percentuais de realização de pronome sujeito preenchido e de sujeito nulo



Fonte: Elaborado pelo autor

Verifica-se, portanto, nos dados em exame, um favorecimento do preenchimento do sujeito em relação ao uso do sujeito nulo. Observamos, então, os resultados de outras pesquisas, como em Duarte (1995, 2019), sobre a frequência do sujeito pronominal nulo no português brasileiro urbano culto e sobre o sujeito em peças de teatro, de Lucchesi (2004, 2009) sobre o português afro-brasileiro, e nossa pesquisa anterior, realizada em 2007, 2021<sup>52</sup>, sobre o sujeito nulo no português popular da Bahia, também realizada na região sudoeste do estado, onde se

<sup>52</sup> Essa pesquisa foi apresentada pelo autor como Dissertação de Mestrado em 2007 e publicada em 2021, sendo citada nas referências deste trabalho. Compararemos com o trabalho atual, pois a pesquisa anterior também foi realizada pelo autor no interior do estado da Bahia em um município que faz limite com o que está sendo estudado. Citaremos, a partir de agora, como Novaes (2007, 2021).

encontra a comunidade quilombola de Lagoinha. No trabalho de Lucchesi (2004, 2009), com uma frequência de 73% de sujeitos realizados e 27% de sujeitos nulos, percebemos uma diferença entre os dois estudos, já que esta última demonstra uma frequência maior, em se tratando da utilização do sujeito preenchido.

Sendo assim, comparamos também com pesquisas realizadas em contextos diferentes, como a de Duarte (1995), em que o resultado também se distancia dos encontrados neste estudo. Duarte (1995) registra 71% de sujeitos preenchidos e 29% de sujeitos nulos. Mesmo que nesta pesquisa se verifiquem valores percentuais inferiores no que diz respeito à variante realização do sujeito pronominal, tal diferença é considerada, de certa forma, previsível. Já em Novaes (2007, 2021), observou-se uma frequência de 58% de pronomes sujeitos realizados e de 42% de sujeitos nulos, o que se aproxima do resultado da pesquisa atual.

A nossa hipótese era a de que, por se tratar do português afro-brasileiro, as taxas gerais de realização do sujeito pronominal superassem a utilização do sujeito nulo. De qualquer forma, os percentuais, apesar de serem menores que os encontrados por Lucchesi (2004, 2009), mostram o favorecimento do sujeito pleno em relação ao sujeito não preenchido. Por isso mesmo, os resultados não deixam de confirmar uma tendência do português brasileiro em relação ao preenchimento do pronome sujeito, haja vista a diferença entre o preenchimento vs. não preenchimento dos dados, ao contrário do que ocorre no português europeu. Sendo assim, em relação aos resultados de Duarte (1995), Lucchesi (2004, 2009) e Novaes (2007, 2021) e na amostra aqui apresentada em nosso trabalho, no que concerne à realização do sujeito pronominal, pode-se verificar que há uma maior tendência de ocorrência do pronome pleno no português brasileiro, sendo que este convive com o sujeito nulo, indicando processo de variação e, provavelmente, de mudança.

A seguir, analisamos as variáveis linguísticas explanatórias, começando com exemplos retirados do *corpus*, para depois explorarmos os resultados.

## 5.2 VARIÁVEIS INDEPENDENTES LINGUÍSTICAS

As variáveis linguísticas explanatórias que trabalhamos abarcam vários níveis da estrutura linguística, desde os morfemas verbais de pessoa e número até o paralelismo discursivo, passando pelo encadeamento das orações, em que o sujeito é referido através de um pronome, realizado ou não na forma fonética da sentença.

Nas seções seguintes, apresentamos os resultados referentes a cada uma das variáveis selecionadas como significativas pelo programa Goldvarb X. Comparamos esses resultados encontrados com o de outros estudos, como os de Lucchesi (2004, 2009d) sobre o português afro-brasileiro, de Duarte (1995, 2019), sobre o português urbano culto e sobre o sujeito em peças de teatro, e de Novaes (2007, 2021), sobre o português popular brasileiro do interior da Bahia em todos os contextos abaixo.

### 5.2.1 Menção anterior do sujeito

Com o grupo de fatores linguísticos *menção anterior ao pronome sujeito*, buscamos avaliar o papel desempenhado pela presença ou ausência de referência prévia ao pronome sujeito. Exibimos, a seguir, exemplos referentes a cada um dos fatores dessa variável, extraídos os dados do nosso *corpus*:

#### **a) presença de um SN na oração anterior:**

(5) O *povo* dançava... Ø dançava. Quando descarregava, Ø tomava a bateria de ôto emprestado. (SL02)

(6) *Finado Vicenço*, nêga, gostava de mim. *Ele* chamava eu era Dôti. (SL06)

#### **b) presença do sujeito pronominal na oração anterior:**

(7) Agora miorô bastante né, que antigamente pa *youê* pegá um carro, unzano aqui atrás, muituzanu aqui atrás, vinte, trinta ano aqui, pa *youê* pegá uma condução aqui, *youê* tinha que ir lá... ir lá pertu de Noguêra (SL03)

(8) Tem o time de futebol, final de semana. A *gente* sai pra Ø jogá torneio tamém. (SL01)

#### **c) presença de sujeito nulo na oração anterior:**

(9) Ø saía cedo. Quando ia chegá lá nas queimada, era quase meio dia, que a *gente* ia caminhano. (SL02)

(10) Ø Disse ó, quando fui lá mais Laurentino, ficava falando que ela soube disso, que ela falô, que foi casado com mãe. (SL05)

#### **d) primeira menção do sujeito:**

(11) Ai, nega, *eu* ando com cabeça véa doida, iguá minino. (SL06)

(12) Hoje... hoje num tem mais, num tem mais isso. Também... tem muitos tempo também. Aquele tempo, *a gente* brincava também de boneca. (SL04)

(13) Aquelas peteca que tinha antigamente. Ø fazia cum pena de galinha e Ø colocava a bolinha e Ø fazia pena de galinha, Ø ficarra brincano tamém, era.

Os resultados relativos a estas variáveis encontram-se dispostos na Tabela 3.

**Tabela 2** - A realização do sujeito pronominal expresso no português afro-brasileiro na comunidade quilombola de Lagoinha – Nova Canaã – BA, segundo a variável menção anterior do sujeito.

Fatores	Aplicação/Total	Percentual	PR
presença de um SN na oração anterior	16/62	25,8	0.230
presença do sujeito pronominal na oração anterior	160/254	63,0	0.486
presença de sujeito nulo na oração anterior	90/254	35,4	0.273
primeira menção do sujeito	359/488	73,6	0.666
Total	625/1058	59,1	

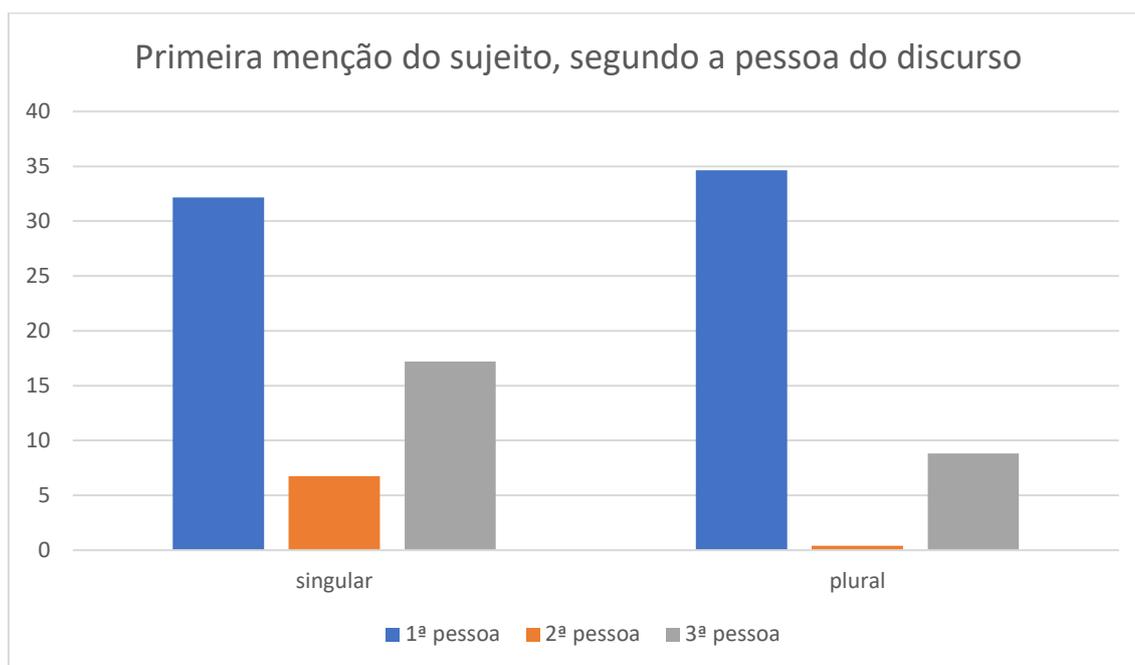
Input = 0.588  
Significance= 0.029

Fonte: Elaborado pelo autor

Conforme exposto na Tabela 2, a primeira menção do sujeito foi o fator que mais favoreceu a ocorrência do sujeito pronominal, com 73,6% de ocorrências e 66 de peso relativo. Por outro lado, a presença de um sintagma nominal na oração anterior foi o fator que mais desfavoreceu o preenchimento do sujeito, com apenas 25,8% de ocorrências de sujeito preenchido e peso relativo de 23, acompanhada das orações antecidas por um sujeito nulo, com 35,4% e 0.27 de peso relativo. A presença do sujeito pronominal na oração anterior também indicou um desfavorecimento da ocorrência do sujeito pronominal com 63% de ocorrência e peso relativo de 48.

Procuramos verificar, então, dentre as 488 ocorrências da primeira menção do sujeito, quais foram às pessoas do discurso que tiveram mais menções. O resultado está exposto abaixo na figura 9.

**Figura 9** - Índices percentuais referentes à primeira menção do sujeito, segundo a pessoa do discurso.



Fonte: Elaborado pelo autor

Percebemos que as primeiras pessoas são as que apresentam os maiores índices, em relação à primeira menção do sujeito, sendo que encontramos a primeira pessoa do plural com 34,63% e a primeira pessoa do singular, com 32,17%. Possivelmente, isso explique o resultado encontrado, que comentaremos a seguir, de como se distribuíram as ocorrências do sujeito pronominal pelas três pessoas do discurso. É possível verificar que a maior ocorrência de sujeitos pronominais preenchidos nas primeiras pessoas tem uma relação com a primeira menção do sujeito, o que pode ser explicado pela necessidade de uma forma mais explícita para a introdução de um novo referente.

Verificando os resultados apresentados com os encontrados na pesquisa realizada por Lucchesi (2004, 2009) sobre o português afro-brasileiro, temos então os seguintes resultados referindo-se à presença do sujeito nulo: 50% de ocorrências e peso relativo de 74, para as orações antecidas por sujeito nulo; 48% de

ocorrências e 63 de peso relativo, para antecedidas por um SN; 25% de ocorrências e 50 de peso relativo nas orações antecedidas por sujeito pronominal realizável; e 21% de ocorrências e 44 de peso relativo, para a primeira referência ou referência isolada. Como podemos perceber as orações antecedidas por sujeito nulo e por um sintagma nominal foram as que mais desfavoreceram a presença do sujeito realizado, o que se aproxima dos nossos resultados.

Foi observado também, em nossos estudos anteriores sobre o português popular do interior do estado da Bahia (NOVAES, 2007, 2021), que a presença do sujeito nulo na oração anterior desfavoreceu bastante a realização do sujeito pronominal, com 12% de ocorrências e peso relativo de 09, seguido da presença do SN, com 10% de ocorrências e peso relativo de 14. A presença do sujeito pronominal indicou um leve desfavorecimento da realização do sujeito pleno, apresentando 53% de ocorrências e um peso relativo de 44. Já a primeira menção do sujeito demonstrou ser um importante fator em relação à realização do sujeito pronominal, apresentando 77% de ocorrências e peso relativo de 69, o que pode ser explicado pela necessidade de uma forma mais explícita para a introdução de um novo referente, como já citado acima.

Vimos, então, que o fator que mais favoreceu o preenchimento do sujeito, tanto em nossa pesquisa, quanto nos estudos de Lucchesi (2004, 2009) e em nosso trabalho anterior (NOVAES, 2007, 2021), foi a primeira menção do sujeito. A ausência do sujeito na oração anterior e a presença de um sintagma nominal foram os fatores que favoreceram o sujeito nulo na oração, tanto em nossa pesquisa, quanto nos outros trabalhos resenhados.

De acordo com Lucchesi (2004), é importante salientar o mecanismo da correferência na ocorrência do sujeito nulo, já que a presença de um SN também favorece a não realização do sujeito pronominal, confirmando a teoria de Chomsky (1988 [1981]), segundo a qual a correferência conjugada à morfologia flexional do verbo seriam os principais licenciadores do sujeito nulo. Vários autores como Galves (1987, 2001), Duarte (1993, 2019), Kato (1999), Barbosa *et al.* (2005) correlacionam o enfraquecimento da flexão em pessoa e número com o aumento de sujeito realizados.

Sendo assim, a perda do valor referencial da flexão colabora para uma menor ocorrência de sujeitos realizados, a exemplo da correferência, onde há um antecedente na oração matriz. Também, de acordo com Almeida (2005), a

identificação é possível, quando há a ausência da flexão, através da ligação do sujeito nulo com um sujeito pronominal correferencial. Sendo assim, é possível identificar o sujeito nulo tanto pela flexão verbal, quanto por meio da correferencialidade, no caso da ausência de marcas morfológicas. De acordo com Duarte (1995), em seus estudos sobre a fala espontânea no Rio de Janeiro, quando há correferência foram encontrados 32% de sujeito nulo, enquanto com o sujeito não referente, somente 12%.

No caso dessa variável, considera-se o conceito de *paralelismo formal* de Naro e Scherre (1993), considerado um aspecto do comportamento linguístico em que há uma tendência na manutenção de um padrão de marcação em orações que estejam na mesma sequência, conservando o mesmo referente. Em seus estudos sobre o sujeito nulo no português afro-brasileiro, Lucchesi (2004, 2009) confirmou que a presença de um sujeito nulo na oração anterior favorece a não realização do sujeito pronominal na oração seguinte. Sendo assim, foi observado que o mecanismo de correferência tem se revelado uma variável importante no condicionamento do não preenchimento do sujeito, já que a presença de um SN favorece a realização do sujeito nulo.

### 5.2.2 Pessoa do discurso

Com a observação dessa variável, procuramos verificar como o nível de realização do sujeito pronominal se distribuiu pelas três pessoas do discurso. Vale destacar que foram encontradas somente duas ocorrências com referência à 2ª pessoa do plural, ou seja, quando o falante se refere a mais de um interlocutor. Além do mais, em todas as vezes em que houve o preenchimento do sujeito pronominal, configurou-se como *knockout*<sup>53</sup> pelo programa computacional. Assim, quando foi feita a segunda rodada, todos os casos de *knockouts* foram excluídos. Vejamos, então, alguns exemplos de ocorrências de pronomes realizados ou não em relação à variável pessoa do sujeito:

---

<sup>53</sup> Idem p. 180

**a) 1ª pessoa do singular:**

(14) Foi quando *eu* fui pra São Paulo (SL01)

(15) Eu trabalhava até essa base de onze hora. Aí  $\emptyset$  ia caminhá quato, cinco quilômetro pa... pa chegá intê essa aula. (SL03)

**b) 2ª pessoa do singular:**

(16) Cê fala uma história do... tem várias história dos meus amigo... (SL03)

(17) E *tu* não sai nem pra dá boa tarde. (SL06)

(18) Aí eu queria, já mãe não quis: “ $\emptyset$  Vai casar com Maria não, *tu* é besta, casar por negócio”. (SL05)

**c) 3ª pessoa do singular:**

(19) Tinha um homi chamado Zé Eduardo, aí, *ele* ensinava a noite (SL03)

(20) Aí agora, *ele* criava uns porquim, *ele* criava uns porquim,  $\emptyset$  criava porco,  $\emptyset$  criava pato. (SL06)

**d) 1ª pessoa do plural:**

(21) *A gente* fica dento de casa... (SL04)

(22) *Nóis* fazia festa, festa mesmo. (SL05)

(23) *Nós* fazia bola de saco,  $\emptyset$  pegava um saquin e  $\emptyset$  enchia com outros saco dentro e  $\emptyset$  fazia as bola pra brincá. (SL01)

**e) 3ª pessoa do plural:**

(24) *Eles* morava lá no sertão. (SL01)

(25) Hoje, as criança de hoje,  $\emptyset$  num sabe. (SL02)

(26) Eu carrego os adulto da noite tamém, que *eles* num tiveram oportunidade

pra estudar também. (SL01)

Os resultados relativos à variável *pessoa do sujeito* são divulgados na Tabela 3.

**Tabela 3** - Variável desinência gramatical no português afro-brasileiro na comunidade quilombola de Lagoinha – Nova Canaã – BA

Fatores	Aplicação/Total	Percentual	PR
Primeira do singular	249/396	62,9	0.547
Segunda do singular	27/41	65,9	0.582
Terceira do singular	88/203	43,3	0.455
Primeira do plural	277/461	60,1	0.519
Terceira do plural	30/96	31,2	0.288
Total	671/1197	56,1	

Input = 0.588  
Significance= 0.029

Fonte: Elaborado pelo autor

No *corpus* analisado, buscando observar a ocorrência do sujeito pronominal na comunidade quilombola de Lagoinha, verificamos que a segunda pessoa do singular foi a que mais favoreceu o preenchimento do pronome sujeito, com um percentual de 65,9% e peso relativo de 58. A primeira pessoa do singular também se revelou favorecedora da variante preenchida, apresentando taxa de 62,9% e peso relativo de 54. O fator que menos favoreceu a presença do sujeito foi a terceira pessoa do plural com 31,2% e peso relativo de 28, seguido pela terceira pessoa do singular com 43,3% e 45 de peso relativo. Em relação à primeira pessoa do plural, encontra-se em um estágio intermediário, com percentual de 60,1% e 51 de peso relativo.

Uma observação importante a ser feita diz respeito a segunda pessoa do singular e à primeira pessoa do plural, em que as formas *tu* e *nós* convivem com o pronome *você* e o sintagma nominal *a gente*. Esse paradigma, de acordo com Duarte (2019), evoluiu de um sistema de seis formas distintas, para um paradigma que apresenta quatro formas verbais, devido ao sintagma nominal *a gente* e o

pronome de tratamento *você*, que se combinam com as formas verbais de 3ª pessoa.

De acordo com Duarte (2019), essa mudança de paradigma já começa na análise dos textos das peças de 1930, que mostram o início da neutralização do *tu* e do *você*, já que se observa uma mistura de tratamento, sendo que alguns autores dessa década abandonam o tratamento com *tu*, optando pelo uso de *você*, o que contribui para que se reduza ainda mais o paradigma flexional. O que já vai prenunciar a forte competição entre o *nós* e *a gente* na segunda metade do século XX, conforme afirma Duarte (2019), em que se encontra o predomínio de *a gente*, para referência à primeira pessoa do plural, e um uso mais modesto de *nós*. Já, em relação ao pronome *tu*, Duarte (2019) observa, nesses estudos de peças de teatro, um retorno no último quartel do século XX, em variação com *você*, o que resulta em três ou quatro formas distintivas. Segundo Duarte (2019, p. 100): “A entrada de *você* e *a gente* constitui, entretanto, uma inegável componente no processo de mudança da gramática brasileira”.

O que observamos, em relação a segunda pessoa do singular, nos resultados encontrados em nossa pesquisa é que há uma grande predominância do pronome *você*, com 81,48% de ocorrências, em relação ao uso do pronome pessoal *tu*, com 18,5% de ocorrências, conforme apresentado na Figura 9. Esses resultados demonstram que a perda da forma verbal de segunda pessoa, ocasionada pelo uso do *você*, combinando com a forma verbal de terceira pessoa, exerceu uma forte influência para o preenchimento do pronome sujeito.

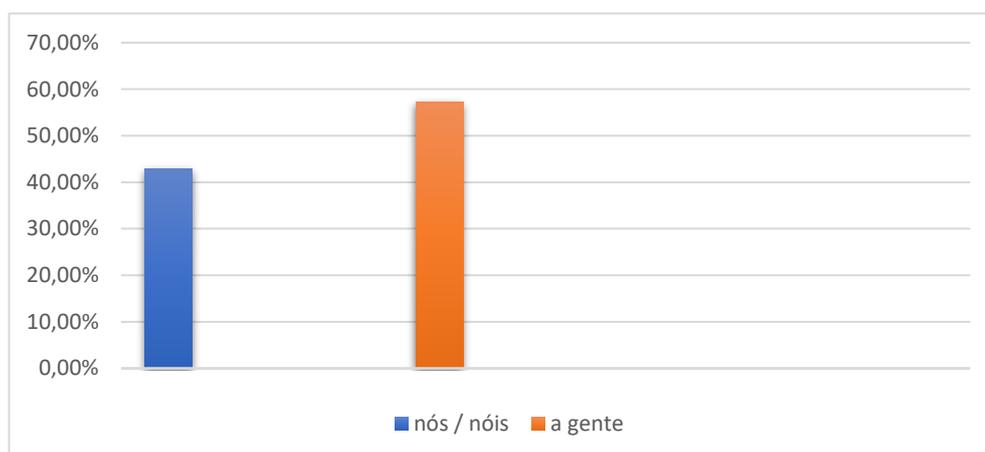
**Figura 10** - Índices percentuais do uso do pronome de 2ª pessoa do singular



Fonte: Elaborado pelo autor

Em relação ao pronome *nós*, encontramos o sintagma nominal *a gente* sendo mais utilizado para referência à primeira pessoa do plural. Foram encontradas 42,9% de ocorrências para o pronome *nós* e 57,0% para o sintagma nominal *a gente*, conforme exposto na Figura 10. Essa utilização do sintagma nominal *a gente*, em preferência ao pronome *nós*, nos permite observar a possibilidade de perda da forma verbal de primeira pessoa do plural, devido ao uso da expressão *a gente*, que combina com a forma verbal de terceira pessoa, o que pode ter favorecido o preenchimento do pronome sujeito.

**Figura 11** - Índices percentuais do uso do pronome de 1ª pessoa do plural



Fonte: Elaborado pelo autor

Observando os resultados acima, podemos levantar algumas questões referentes ao maior preenchimento do pronome sujeito na segunda pessoa do singular, com um percentual de 65,9% e peso relativo de 58. Encontramos o uso do pronome de tratamento  *você*, em substituição ao pronome pessoal *tu*, em que são verificadas 81,48% vs. 18,5% de ocorrências, respectivamente, o que compromete a desinência verbal de segunda pessoa do singular, ao aproximá-la da desinência de terceira pessoa.

Desse modo, o uso do pronome se faz necessário para a diferenciação da pessoa do discurso. Já, em relação a primeira pessoa do plural, em que foi encontrado um percentual 60,1% e 51 de peso relativo, há uma situação de competição entre o pronome *nós* e o sintagma nominal *a gente*, observamos 42,9% e 57,03% de ocorrências respectivamente. Assim, conforme afirma Duarte (2019, p.

101): “é inegável que a plena pronominalização de *você* e *a gente*, combinados com a forma verbal na terceira pessoa do singular, constitui o mais significativo desencadeador da redução do paradigma flexional”.

Verificamos, então, os resultados encontrados por outros autores a respeito do preenchimento ou não do pronome sujeito. Nos estudos realizados por Lucchesi (2004, 2009d), sobre o português afro-brasileiro, a 2ª pessoa do singular mostra-se um fator importante favorecendo a presença do pronome sujeito, assim como no presente estudo. Nesse caso, temos um favorecimento do preenchimento do sujeito pronominal, com sujeito nulo de 17% e 43 de peso relativo. Já a 3ª pessoa do singular é o fator que menos favorece a realização do sujeito pronominal, com 34% de sujeito nulo e 55 de peso relativo. Quanto à 1ª pessoa do singular, verifica-se que está em um estágio intermediário, em relação a essa realização ou não. No entanto, nesses mesmos estudos, a 1ª pessoa do plural é indicada como a que mais favorece essa realização, onde são encontradas 27% de ocorrências de sujeito nulo e peso relativo de 34. Também nos resultados obtidos nos estudos realizados por Duarte (1995) sobre o português brasileiro urbano culto, é encontrada a 2ª pessoa como aquela que mais favorece a realização do sujeito pronominal e a 3ª pessoa como o fator que menos favorece essa realização. Nesse mesmo estudo, encontra-se a 1ª pessoa como um fator em estágio intermediário, em relação à realização pronominal.

Nos resultados relatados em nosso trabalho anterior sobre o sujeito nulo no português popular da Bahia (NOVAES, 2007, 2021), também ficou evidente que a 2ª pessoa do singular é aquela em que o pronome sujeito é mais realizado, com uma frequência de 80% do total de ocorrências, contra 58% de frequência geral, e um peso relativo de 66. Tal fato se deve a motivações discursivas e, principalmente, em função de o uso do pronome *você*, que se conjuga com a forma verbal não marcada da 3ª pessoa do singular, dificultar a interpretação da categoria vazia por parte do destinatário neste contexto. Isso se aplica também ao uso do pronome *tu*, que também é usado na comunidade de fala analisada, mas que se conjuga também com a forma verbal não marcada da 3ª pessoa do singular. Por outro lado, a terceira pessoa, que apresenta uma simetria notável entre o singular e o plural, é aquela que mais desfavorece a realização do sujeito pronominal. No singular, apenas 44% de realização do sujeito pronominal, com peso relativo de 37; no plural, são 50% do total de ocorrências, com peso relativo de 42. Possivelmente, a opção pelo uso do

sujeito nulo na terceira pessoa, ocorre porque o falante assume o morfema zero como 3ª pessoa, havendo então a necessidade de marcar a 2ª e 1ª do plural, com o preenchimento do pronome sujeito. Já a 1ª pessoa do singular, revelou estar em um estágio intermediário, em relação à realização do sujeito pronominal, com um ligeiro aumento de 58 para 60% do total de ocorrências, de modo que se obteve um peso relativo de 51, praticamente na marca da neutralidade. Entretanto, a 1ª pessoa do plural favoreceu a realização do sujeito pronominal, com 65% do total de ocorrências, com um peso relativo 61. Esse resultado pode ser atribuído ao elevado uso da expressão nominal *a gente*, gramaticalizada como pronome de 1ª pessoa do plural. Essa forma normalmente se combina com a forma verbal não marcada da 3ª pessoa do singular. Além disso, é frequente a combinação do pronome pessoal *nós* com essa forma verbal não marcada.

Vejamos a seguir a análise referente ao traço semântico do sujeito, que tem se revelado de grande importância no estudo do preenchimento ou não do pronome sujeito no português do Brasil.

### 5.2.3 Traço semântico do sujeito

Ao examinarmos a *categorização semântica do sujeito*, foram considerados os seguintes fatores os quais ilustramos com alguns exemplos encontrados em nosso *corpus* de pesquisa:

#### **a) [+ animado] / [+ humano]**

(27) Ixi, *nóis* jogando bola lá no campo lá, os minino, aí chegô uma vaca lá. (SL01)

(28) Ø Trabalhava nessa roça de arroz. (SL02)

#### **b) [- animado] / [- humano]**

(29) Passava no brejo, derrubava *a madeira* pro lado de lá, *ela* tombava. (SL05)

(30) *Chão* que não tinha cimento, Ø tinha nada. (SL02).

**c) [+ animado] / [-humano]**

(31) aí chegô *uma vaca* lá e nós foi tirá ela pra fora da manga lá do campo...

Aí *ela* correu atrás de nós. (SL01)

(32) Aí ele foi pelejando com *essas ovêa*. Aí Ø não rendeu muito não, Ø num deu muita sorte não.

Na Tabela 4, a seguir, exibimos os resultados referentes a este grupo de fatores.

**Tabela 4** - Variável traço semântico do sujeito no português afro-brasileiro na comunidade quilombola de Lagoinha – Nova Canaã – BA

Fatores	Aplicação/Total	Percentual	PR
+ animado / - humano	3/10	30,0	0.273
+ animado / + humano	665/1159	57,4	0.513
- animado / - humano	5/30	16,7	0.154
Total	673/1199	56,1	

Input = 0.588  
Significance= 0.029

Fonte: Elaborado pelo autor

O resultado dessa variável nos possibilitou observar que o fator [+ animado / + humano], apresentou 57,4% de ocorrências e 51 de peso relativo, em relação ao preenchimento do pronome sujeito. Já o fator [- animado / - humano], por outro lado, exibiu um forte desfavorecimento do preenchimento do pronome sujeito, sendo encontrados 16,7% de ocorrências e peso relativo de 15, favorecendo assim a ocorrência do sujeito nulo.

Desse modo, podemos observar a importância do traço [+ animado] em relação ao preenchimento do sujeito. Ao observarmos os estudos de Duarte (1995, 2019), que afirmou ser essa variável bastante relevante em relação à implementação de categorias vazias no português do Brasil, conforme veremos adiante, e os estudos de Lucchesi (2004, 2009d), percebemos que esta pesquisa se aproxima dos resultados apresentados pelos autores.

Nos estudos de Duarte (1995), observou-se a ocorrência de 32% para o traço [+ animado] e de 44% para o traço [- animado] em relação à utilização do sujeito nulo. Tais resultados são, em certa medida, confirmados também por Lucchesi (2004, 2009d), sobre o português afro-brasileiro, no qual foram encontradas 63% de ocorrências e peso relativo de 73 para o traço [- animado] e 29% de ocorrências e peso 47 para o traço [+ animado] no uso do sujeito pronominal nulo, o que favorece também o preenchimento do sujeito. Isso foi observado por Duarte (1989) ao estudar a representação do objeto anafórico. Esses estudos revelaram também a importância do traço [+/- animado]: foram encontrados 23,7% para o traço [+ animado] e 76,3% para o traço [- animado] em relação ao apagamento do objeto.

Em nossa pesquisa sobre o português popular do Brasil (NOVAES, 2021, 2007), os resultados dessa variável indicaram que o traço semântico [- humano] favoreceu consideravelmente a não realização do sujeito de 3ª pessoa, em que foi encontrada uma frequência de 17% de sujeitos nulos, com peso relativo de 17. Já o traço semântico [+ humano], com 59% de ocorrências e 51 de peso relativo, demonstrou ser um fator de favorecimento em relação à realização do sujeito pronominal. Esses estudos também confirmaram que esse traço é um dos mais relevantes em relação à categoria vazia no português popular do Brasil.

O traço semântico [+ humano /- humano] tem se revelado um dos fatores muito relevantes, conforme a afirmação de Duarte (1989, 1995, 2019), o que também é afirmado por Lucchesi (2004, 2009d), para a realização ou não do sujeito pronominal. Duarte (1989, 1995, 2019) afirma ainda que essa é uma das variáveis que mais interfere não só em relação à realização ou não do sujeito, mas também em se tratando da realização ou não do objeto nulo. Enquanto o traço semântico [+ humano] favorece a realização do sujeito pronominal e do objeto anafórico, o traço semântico [- humano] favorece o sujeito e o objeto nulos. Segundo Duarte (1995), podemos concluir então que o traço [- humano] é um fator de grande importância quando estão em jogo as categorias vazias, tendo-se revelado um dos fatores responsáveis pela origem e implementação, no português do Brasil, do objeto nulo.

Isso vem ocorrendo também em relação ao sujeito, quando se leva em consideração a perda gradual da propriedade *pro-drop*, a qual foi observado por Lucchesi (2004, 2009d) em relação ao traço [- humano], que favorece consideravelmente a realização do sujeito nulo de 3ª pessoa. Outro fator de natureza semântica e que também se apresenta como condicionador de importância, de

acordo com Duarte (2019), é o [+/- animado]. Segundo a autora, o traço [+ animado] se mostra como favorecedor do preenchimento do sujeito, o que corrobora a hierarquia referencial proposta em Cyrino, Duarte & Kato (2000) para a realização de pronomes nulos e plenos em posição de sujeito e objeto.

Analisaremos, a seguir, a desinência gramatical do verbo.

#### 5.2.4 Desinência gramatical do verbo

Nesta análise<sup>54</sup>, o morfema zero foi escolhido como um fator de grande importância, pois pode servir a várias pessoas, conforme exemplos expostos anteriormente, enquanto o morfema de 1ª pessoa do singular só é realizado em alguns tempos verbais. O morfema de 1ª pessoa do singular é o único que não é ambíguo, pois se refere somente a essa pessoa, não sendo utilizado por nenhuma das outras pessoas do discurso. É o morfema mais informativo da pessoa do sujeito. O mesmo não acontece com o morfema de 3ª pessoa do plural que tanto pode se ligar à 2ª quanto à 3ª pessoa do plural. Quanto ao morfema de 1ª pessoa do plural, foram encontrados 4 dados, enquanto, em relação à 3ª pessoa do plural, foi encontrado apenas 1 dado referente a este morfema, o que foi configurado como *knockout* pelo Golbdvarb X. Desse modo, só apresentamos na Tabela 6 os resultados relacionados ao morfema zero e morfema de 1ª pessoa do singular.

Ao estudarmos essa variável, foram levados em conta os seguintes fatores:

##### **a) morfema zero<sup>55</sup>**

(33) Desde dez ano mais ou menos, oito ano em diante, *eu já trabalhava* com meu pai já na... na roça. (SL03)

---

<sup>54</sup> Apresentaremos aqui um cômputo geral, que será detalhado mais adiante com o cruzamento da pessoa do discurso com a desinência gramatical do verbo.

<sup>55</sup> O morfema zero é constituído pela ausência de uma marca de oposição, relacionada a outro termo marcado, que possui a mesma função ou significação do morfema presente, correspondente à sua posição. Sendo assim: “A ocorrência do morfema zero na flexão verbal se dá pela falta de marca em relação às outras do mesmo paradigma. Em: escrevo, escreves, escreve, escrevemos, escreveis, escrevem, a 3ª forma verbal é caracterizada pela falta de desinência em relação às outras pessoas verbais. Temos, portanto, um morfema zero”. (MACHADO, 2002, p. 7).

(34) Vicenço, vê se *tu encontra* lá. (SL06)

(35) Na estrada, se *you* *viaja*, quando *you* *tá* em casa,  $\emptyset$  *tá* sempre com medo. (SL04)

(36) *Ele vai* ver qual é os pessoá, todo mundo bem sabe disso.

(37) *Nóis saía* lá em cima na ponte que vai pra Morrinhos. (SL05)

(38) *A gente...* *fazia* tipo de comida,  $\emptyset$  *buscava* folha nas roça,  $\emptyset$  *buscava* as serralha. (SL02)

(39) Corre minha fia, vai lá pá dento do quarto, pá ela num vê *you* *cumê* esse pirão d'água

(40) Não, *eles falava* que era... Isqueci o nome que *eles falava* que era... eu acho que era difunto, essas coisa assim (SL01).

#### b) morfema de 1ª pessoa singular

(41) *Eu não quero* desisti. (SL04)

(42)  $\emptyset$  Num *tenho* muita lembrança das história dela. (SL01)

**Tabela 5** - Variável sexo no português afro-brasileiro na comunidade quilombola de Lagoinha – Nova Canaã – BA

Fatores	Aplicação/Total	Percentual	PR
Morfema zero	570/1054	54,1	0.485
Morfema de 1ª p. singular	102/140	72,9	0.610
Total	672/1194	56,3	

Input = 0.588

Significance= 0.029

Fonte: Elaborado pelo autor

Nesta análise, o morfema zero foi escolhido como um fator de grande importância, pois pode servir a várias pessoas, conforme exemplos expostos anteriormente, enquanto o morfema de 1ª pessoa do singular, só é realizado em

alguns tempos verbais. O morfema de 1ª pessoa do singular é o único que não é ambíguo, pois se refere somente a essa pessoa, não sendo utilizado por nenhuma das outras pessoas do discurso. É o morfema mais informativo da pessoa do sujeito. O mesmo não acontece com o morfema de 3ª pessoa do plural que tanto pode se ligar à 2ª quanto à 3ª pessoa do plural. De acordo com os resultados apresentados na Tabela 6, verificamos que o morfema de 1ª pessoa do singular, com 72,9% de ocorrências e peso relativo de 61, foi o que mais favoreceu a presença do pronome sujeito. Sendo assim, o morfema de 1ª pessoa do singular demonstra maior significância em relação à presença do pronome sujeito. Quanto ao morfema zero, verificamos um leve desfavorecimento, em relação ao preenchimento, ao apresentar 54% e 48 de peso relativo.

Ao observarmos os estudos de Lucchesi (2004, 2009d) sobre o português afro-brasileiro, verificamos que o morfema de 1ª pessoa do singular encontra-se em um estágio intermediário, sendo contabilizadas 29% de ocorrências correspondentes a um peso relativo de 51, em relação ao uso do sujeito pronominal nulo. Nos estudos de Novaes (2007, 2021), encontramos, em relação ao morfema de 1ª pessoa do singular, 62% de ocorrências e 50 de peso relativo em relação ao preenchimento do pronome sujeito, o que também indica um estágio intermediário. Por outro lado, em nossos estudos, o morfema de 1ª pessoa do singular demonstrou um favorecimento ao preenchimento do pronome sujeito, apresentando 72,9% de ocorrências e 61 de peso relativo.

Quanto ao morfema zero, verificamos nos estudos realizados por Lucchesi (2004, 2009d) sobre o português afro-brasileiro, em relação ao uso do sujeito pronominal nulo, que o morfema zero apresentou 26% de ocorrências e peso relativo de 48, indicando uma pequena queda, que favorece o uso do sujeito pronominal pleno. Já em nossos estudos sobre o português popular brasileiro, em Novaes (2007, 2021), foram encontradas 57% de ocorrências e 51 de peso relativo, um estágio intermediário em relação ao preenchimento. Tais resultados se aproximam também da nossa pesquisa, como podemos ver acima.

Essa conclusão sobre o licenciamento do sujeito nulo também é apresentada por Duarte (2019), ao afirmar que essa ocorrência pela perda do traço semântico de pessoa na categoria gramatical é responsável pela identificação das três pessoas gramaticais. Desse modo, temos, então, o traço sintático que nos leva aos quatro morfemas analisados em nossos estudos: *-o*, *-ei* (1ª pessoa do plural); *-mos* (1ª

pessoa do plural); *-m* (3ª pessoa do plural) e morfema zero. O que se tem observado, conforme afirma Duarte (2019), é que a desinência *-mos* está deixando de ser utilizada e as desinências *-o* e *-ei* estão limitadas a dois tempos verbais (presente, pretérito perfeito do indicativo). A desinência *-m*, por seu turno, tem sido omitida, cada vez com maior frequência, nos registros mais informais, e a sua omissão, principalmente quando se trata de falantes com uma baixa escolaridade. Essa situação contribui para que a desinência zero seja marcada negativamente em relação ao número e pessoa, mostrando a perda da flexão isoladamente. Por isso mesmo, confirmando os achados de Duarte (2019) e Lobato (1986), fica mais uma vez comprovada a importância da flexão verbal no licenciamento do sujeito nulo e no preenchimento do pronome sujeito.

Em se tratando das línguas românicas, Chomsky (1988 [1981]) também classifica essa variável como um dos principais aspectos no que se refere ao licenciamento do sujeito nulo. Por outro lado, temos ainda a afirmação de Lobato (1986) de que nas línguas *pro-drop*, a flexão verbal funciona como principal regente para a posição de sujeito, por permitir a utilização do sujeito nulo, ao se relacionarem com a flexão manifesta. Duarte (1995), ao afirmar que o português brasileiro passa por um processo de importantes mudanças em relação à realização do pronome sujeito, também afirma que esse processo está relacionado à simplificação do modelo flexional, devido à redução ocorrida no quadro pronominal. Sobre o morfema de 3ª pessoa, Duarte (2019) comenta que, apesar da 1ª e 2ª pessoas estarem perdendo a opção pelo sujeito nulo, o mesmo não acontece com a 3ª pessoa que continua a usufruir dessa opção, o que se deve à perda da riqueza funcional do paradigma.

#### 5.2.5 Cruzamento da pessoa do discurso com a desinência gramatical do verbo

Comparando os resultados acima, em relação ao preenchimento do pronome sujeito, referentes à pessoa do sujeito e à desinência gramatical do verbo, percebemos, então, a necessidade da realização de um cruzamento dos dados, utilizando o Goldvarb X, para que pudéssemos fazer uma análise mais apurada, o que pode ser verificado na Tabela 7 abaixo.

Vale salientar que, em relação à 2ª e à 3ª pessoa do singular, acompanhadas de seus respectivos morfemas, não foram consideradas porque, no PB, não existe a

possibilidade de virem acompanhadas de qualquer morfema flexional. Também só houve uma ocorrência da 2ª pessoa do plural com o pronome *vocês*, relacionado a esse morfema flexional. O mesmo aconteceu com o pronome *nós*, acompanhado de sua desinência em que foram encontrados somente quatro dados, o que, nesses casos, foi considerado como *knouckout* pelo programa computacional.

**Tabela 6** - Variável escolaridade no português afro-brasileiro na comunidade quilombola de Lagoinha – Nova Canaã – BA

Fatores	Aplicação/Total	Percentual	PR
1ª pessoa singular com morfema zero	148/257	57,6%	0.529
2ª pessoa singular com morfema zero	27/41	65,9%	0.560
3ª pessoa singular com morfema zero	88/203	43,3%	0.456
1ª pessoa plural com morfema zero	277/458	60,5%	0.501
3ª pessoa plural com morfema zero	28/94	29,8%	0.275
1ª pessoa do singular com morfema de 1ª pessoa do singular	101/139	72,7%	0.649
Total	669/1192	56,1	

Input = 0.574  
Significance= 0.007

Fonte: Elaborado pelo autor

Ao realizarmos o cruzamento das variáveis pessoa do discurso vs. a desinência gramatical do verbo, observamos em relação ao morfema zero, que o maior preenchimento do pronome sujeito ocorre na 2ª pessoa do singular com morfema zero, que apresenta 65,9% de ocorrências e peso relativo de 56. Em seguida temos a 1ª pessoa do singular e a 1ª pessoa do plural, associadas ao morfema zero, com 57,6% e 60,5% de ocorrências, e 52 e 50 de peso relativo, respectivamente. Quanto à 3ª pessoa do singular, com o morfema zero, apresentou 43,3% de ocorrências e peso relativo de 45. A 3ª pessoa do plural, relacionada ao morfema zero, foi a que menos favoreceu o preenchimento do pronome sujeito com 29,8% e 27 de peso relativo. Quando ao cruzamento, em relação à desinência própria de cada pessoa, observamos que o preenchimento do pronome sujeito ocorreu na 1ª pessoa do singular com o morfema de 1ª pessoa do singular, com 72,7% de ocorrências e 64 de peso relativo. Possivelmente, isso se deve ao fato de que não foram encontrados dados suficientes, conforme citado acima, sendo

considerado como *knockout* pelo programa computacional, já que houve somente uma ocorrência da 2ª pessoa do plural com o pronome *vocês* e quatro dados referentes o pronome *nós*, acompanhado de sua desinência específica. De acordo com o que verificamos nas tabelas anteriores, em relação ao preenchimento do pronome sujeito nas variáveis: pessoa do discurso e desinência gramatical do verbo, os resultados foram muito próximos ao cruzamento aqui apresentado.

A partir da análise da relação entre o preenchimento do pronome sujeito e a presença de um morfema flexional, é possível verificar com maior nitidez a presença ou não do pronome pleno, a partir do cruzamento dessas duas variáveis analisadas até então. Ao verificarmos a tabela 7, que demonstra a frequência de cada pessoa do discurso, levando-se em conta a presença ou ausência de seu morfema flexional próprio, é possível percebermos a existência da relação entre a ausência de uma marca flexional e a realização do sujeito pronominal. Sendo assim, temos a 2ª pessoa do singular relacionada ao morfema zero, em que a ausência total do morfema favorece a utilização do sujeito pronominal, conforme já citado anteriormente. Já as primeiras pessoas, tanto do singular, quanto do plural, associadas ao morfema zero, demonstraram estar em um estágio intermediário, sendo que a 1ª do singular indica um leve favorecimento do pronome sujeito.

Quanto à 1ª pessoa do plural, provavelmente a explicação estaria, possivelmente na saliência fônica desse morfema, conforme indica Lucchesi (2009d), e devido à competição entre o pronome *nós* e a expressão nominal *a gente*, conforme afirma Duarte (2019). Já em relação à 1ª pessoa do singular, possivelmente há uma forte influência da 1ª pessoa do plural, pois o mesmo pode ser observado em relação à pessoa do discurso, em que os resultados são bem próximo entre essas duas pessoas. Ao relacionarmos a 3ª pessoa do singular com o morfema zero, percebemos um leve desfavorecimento do preenchimento do pronome sujeito, sendo que a 3ª pessoa foi a que mais desfavoreceu o uso do pronome pleno. Essa opção pelo uso do sujeito nulo na terceira pessoa ocorre, provavelmente, porque o falante assume o morfema zero como 3ª pessoa, havendo então a necessidade de marcar a 2ª e 1ª do plural, com o preenchimento do pronome sujeito, conforme já citado acima.

Temos então, por outro lado, um maior favorecimento do pronome sujeito com a 1ª pessoa do singular associada ao morfema de 1ª pessoa do singular, o que muito nos surpreendeu. Essa relação apresenta uma maior significância com a 1ª

pessoa do singular, que, mesmo com a presença do morfema de 1ª pessoa, o que por si só já definiria a pessoa do discurso, demonstra um maior favorecimento do preenchimento do pronome sujeito, não implicando em um aumento significativo na frequência de uso do sujeito nulo. Segundo Lucchesi (2009d), esse fato pode ser decorrente de o indivíduo estar se referindo a si mesmo, devido a fatores de ordem pragmática e/ou ideológica, o que pode estar intervindo nesse caso. Essa hipótese merece ser verificada em futuras investigações, em um maior aprofundamento sobre essa questão.

Ao verificarmos os resultados apresentados acima, observamos que, em relação à segunda pessoa do singular, a ausência total da desinência gramatical favoreceu o uso do pronome sujeito. Nesse caso, encontramos 65,9% de ocorrências do pronome sujeito, em relação à variável pessoa do discurso, o que confirma a influência da mudança de paradigma, de acordo com Duarte (2019), com a substituição do uso do pronome pessoal *tu* pelo pronome de tratamento *você*, já que foi verificada uma grande predominância do pronome de tratamento *você*, com 81,48% de ocorrências, enquanto o uso do pronome pessoal *tu*, revelou 18,5% de ocorrências. Em relação à segunda pessoa, segundo Duarte (2019, p. 108), “o processo de mudança no PB está praticamente concluído”.

A mesma situação, relacionada à perda do paradigma citado por Duarte (2019), ocorre na primeira pessoa do plural, em que a substituição do pronome pessoal *nós*, pelo sintagma nominal *a gente* também favorece o preenchimento do sujeito pronominal. Como vimos acima na figura 10, há uma situação em que o sintagma nominal *a gente*, apresenta 57,03% de ocorrências, enquanto o pronome pessoal *nós*, demonstra 42,9% de ocorrências para referência à primeira pessoa do plural. Verificamos, então, de acordo com a Figura 11, apresentada acima, que a variável pessoa do discurso apresenta 60% de ocorrências.

Já em relação à variável desinência gramatical do verbo, foram poucos os dados encontrados referentes à desinência de primeira pessoa do plural, sendo apenas 4, o que se configurou como *knockout* pelo programa computacional. Em relação a esse contexto, de acordo com a afirmação de Duarte (2019, p. 97), “é inegável que a plena pronominalização do *você* e *a gente*, combinados com a forma verbal na terceira pessoa do singular, constitui o mais significativo desencadeador da redução do paradigma flexional”. Por isso mesmo, essa ocorrência da presença do *você* e *a gente*, conforme o autor pode ser constituída como, sem sombra de

dúvidas, uma componente de mudança da gramática brasileira. Já na primeira pessoa do singular, foram encontrados 56,3% de ocorrências do sujeito preenchido, em relação à pessoa do discurso, 61,8% de ocorrências em relação à desinência gramatical. Esse resultado demonstra que, provavelmente, por influência do contexto apresentado na primeira pessoa do plural, tenha ocorrido o favorecimento do uso do sujeito pronominal pleno, levando-se em conta, ainda que, conforme Duarte (2019), a 1ª pessoa apresenta-se como um fator neutro em relação à realização pronominal.

Analisando a terceira pessoa do singular, encontramos 43,3% de ocorrências em relação à pessoa do discurso e 54% em relação à desinência gramatical (morfema zero), o que pode ser explicado, conforme já citado anteriormente, possivelmente, porque o falante assume o morfema zero como 3ª pessoa. Diante desse contexto, há um desfavorecimento da realização do sujeito pronominal, já que ocorre a necessidade de marcar a 2ª e a 1ª do plural, com o preenchimento do pronome sujeito. No caso específico da terceira pessoa do plural, em que ocorreram 31,2% de ocorrências relacionadas à pessoa do discurso, em relação à desinência, deve-se levar em conta também os poucos dados encontrados com a presença do morfema referente a esta pessoa do discurso, apenas 1, o que se configurou como *knockout*, o que também deve ser ressaltado em relação à primeira pessoa do plural. A omissão da desinência *-m*, segundo Duarte (2019), como já citado anteriormente, tem sido cada vez mais frequente quando se trata de registros informais e quando omitida, principalmente, em relação aos falantes de baixa escolaridade.

Como destacado pelos autores mencionados, a realização ou não do sujeito pronominal em cada pessoa do discurso está fortemente relacionado à presença de um morfema específico de pessoa e número na forma verbal que acompanha esse sujeito pronominal. Por isso, é imperioso focalizar diretamente a presença desse morfema específico de pessoa e número como uma variável explanatória da realização do sujeito pronominal.

### 5.2.6 Conclusão sobre as variáveis linguísticas

Do total de 1199 ocorrências computadas, 673 (56,1%) apresentam o sujeito pronominal pleno, enquanto 526 (43,9%) o sujeito nulo. Analisando os resultados relativos ao preenchimento *versus* a ausência do pronome sujeito, verificamos um favorecimento do pronome sujeito em nossa amostra, o que indica que há mais ocorrência do pronome pleno em relação ao sujeito nulo no português afro-brasileiro, ao contrário do que ocorre no português europeu.

Seguindo a ordem de significância considerada pelo Goldvarb X, durante a análise multivariada, temos na variável *menção anterior do sujeito* uma maior realização do sujeito pronominal no fator primeira menção do sujeito, em que foi encontrado um peso relativo de 66. Quanto à *pessoa do discurso*, observou-se que o sujeito pronominal é mais realizado na 2ª pessoa do singular, que apresentou 58 de peso relativo. Isso se deve à substituição do pronome *tu* por *você* e à perda da forma marcada de segunda pessoa, tornando necessária a realização do pronome sujeito para que se possa identificar a pessoa do discurso. Em seguida, temos a 1ª pessoa do singular com 54 de peso relativo. Já a 1ª pessoa do plural, encontra-se em um estágio intermediário com 51 de peso relativo, onde pode ser observada a substituição do pronome *nós* pela expressão *a gente*, o que também retira a forma verbal marcada da primeira pessoa do plural.

Os estudos têm revelado que a primeira e a segunda pessoa estão mais avançadas na mudança e isso é explicado pela hierarquia referencial mostrada por Cyrino Duarte e Kato (2000). A terceira pessoa tem revelado mais resistência, porque ao contrário da primeira e da segunda, são dêiticas. A terceira é anafórica, tem um antecedente que a identifica. Essa perda da “riqueza funcional” não permite mais identificar o português brasileiro como uma língua *pro-drop*. Já em relação ao traço semântico do sujeito, o fator [+ animado / + humano] demonstra um estágio intermediário em relação ao preenchimento do pronome sujeito, com 51 de peso relativo. No entanto, indica uma grande diferença em relação ao fator [- animado / - humano] que apresentou 15 de peso relativo, o que confirma as afirmações de Duarte (1989, 1995, 2019) e Lucchesi (2004, 2009d) de que o traço semântico tem se revelado um dos fatores muito relevantes para a realização ou não do sujeito pronominal.

Percebemos que os fatores menção anterior do sujeito, a 2ª pessoa do discurso e a desinência verbal de 1ª pessoa do singular foram os que mais favoreceram o preenchimento do sujeito nos dados encontrados no *corpus* de nossa pesquisa do português popular afro-brasileiro da comunidade quilombola de Lagoinha. Isso nos leva a concluir que a morfologia verbal e a correferencialidade foram importantes para que houvesse a ocorrência do pronome sujeito nas orações.

Esses resultados sugerem a necessidade de uma investigação maior a respeito do Sujeito Nulo, que é um dos parâmetros da Gramática Universal. Por outro lado, em relação ao português brasileiro nos chama a atenção para uma mudança de paradigma sobre a sua classificação em se tratando de língua *pro-drop*. O mecanismo de correferência parece estar se mostrando bastante influente em relação ao preenchimento do pronome sujeito, no entanto, esses estudos indicam que são necessários mais aprofundamentos sobre os mecanismos de mudança no português brasileiro, já que por si só não são conclusivos.

A seguir apresentaremos os resultados da análise quantitativa das variáveis sociais na realização ou não do sujeito pronominal.

### 5.3 VARIÁVEIS INDEPENDENTES EXTRALINGUÍSTICAS

Uma das características universais das línguas naturais é a variação linguística como já citada anteriormente, que convive com efeito de agentes internos e externos ao ambiente linguístico. De acordo com Mollica (2004), apesar de apresentar-se aparentemente caótica e aleatória, essa heterogeneidade da língua é regular, sistemática e previsível e, por isso mesmo, a Sociolinguística tem esse elemento como um dos pontos principais, ao analisar a variação linguística e as mudanças que possam vir a ocorrer, por meio, principalmente, da observação de fatores extralinguísticos. Sendo assim: “*La lengua es una forma de conducta social; como tal, creada y moldeada por seres humanos*”<sup>56</sup> (SILVA-CORVALÁN, 1988, p. 68).

Para verificar a relação da língua com a sociedade, buscamos observar a influência na realização ou não do sujeito pronominal das seguintes variáveis sociais: sexo, idade, grau de escolaridade e estada fora da comunidade. No entanto,

---

<sup>56</sup> A língua é uma forma de conduta social; como tal, criada e moldada por seres humanos (tradução nossa).

como já foi dito anteriormente, somente a variável sexo foi considerada relevante pela análise multivariada, não sendo atribuída significância às demais variáveis sociais. Isso se deve ao fato de os informantes possuírem praticamente o mesmo grau de escolaridade, sendo todos eles analfabetos ou semianalfabetos. Quanto à estada fora da comunidade, poucos foram os informantes que tinham morado em outros lugares e, quando o fizeram, foi por um curto período. Não levamos em consideração a localização dos informantes, pois todos os entrevistados são moradores da zona rural, na comunidade quilombola de Lagoinha.

Em relação às variáveis sociais, pode-se perceber que não houve uma diferença significativa entre os fatores utilizados para a análise dessas variáveis, apesar de percebermos uma diferença no uso do pronome sujeito entre os falantes do sexo masculino e feminino. O resultado, veremos a seguir na Tabela 08, em relação à variável que foi considerada como significante pelo *Goldvarb X*.

### 5.3.1 Sexo

Ao analisarmos a variável sexo, levamos em conta que os papéis exercidos pelos homens e mulheres na sociedade devem ser considerados fundamentais, já que os estudos sociolinguísticos, muitas vezes, demonstraram diferenças na aplicação de fenômenos variáveis a depender do sexo dos indivíduos. Buscando verificar resultados que comprovassem essas afirmações, sobre a variável sexo, foram utilizados os seguintes fatores: masculino; feminino.

Os resultados referentes a essa variável podem ser observados na Tabela 08:

**Tabela 7-** A realização do sujeito pronominal expresso no português afro-brasileiro na comunidade quilombola de Lagoinha Nova Canaã – BA, segundo o cruzamento pessoa do discurso vs. desinência verbal

Fatores	Aplicação/Total	Percentual	PR
Masculino	327/682	47,9	0.412
Feminino	346/517	66,9	0.615
Total	673/1199	56,1	

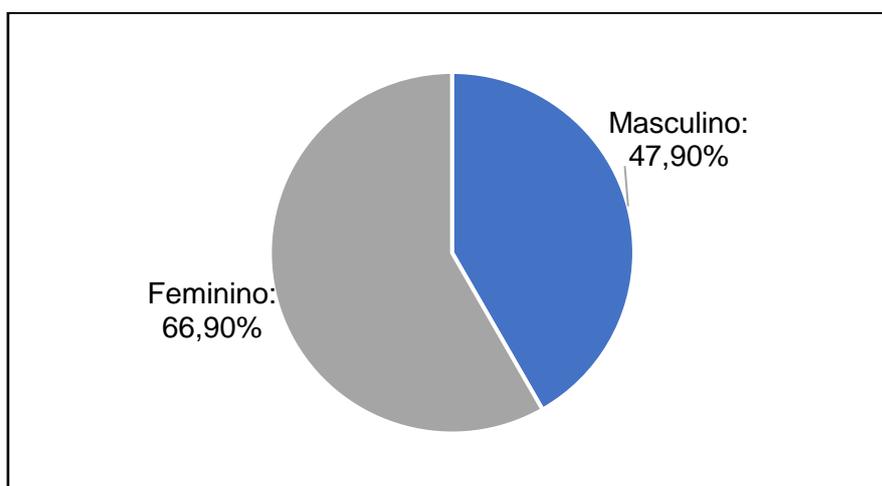
Input = 0.588

Significance= 0.029

Fonte: Elaborada pelo autor

Encontramos, em nossa pesquisa, de acordo com os resultados apresentados na Tabela 8, em relação à variável sexo, que falantes do sexo feminino apresentam 66,9% de ocorrências do pronome sujeito e 61 de peso relativo, enquanto o falante do sexo masculino apresentou 47,9% de ocorrências e peso relativo de 41. Percebe-se, desse modo, que o uso do sujeito pronominal foi mais realizado pelos falantes do sexo feminino do que entre os falantes do sexo masculino, conforme verificamos no gráfico abaixo, apresentado na figura 11, em que podem ser visualizados esses resultados.

**Figura 12** - Índices percentuais de realização de pronome sujeito preenchidos e de sujeito nulo de acordo com o sexo.



Fonte: Elaborado pelo autor

Ao verificarmos esses resultados, observamos os estudos sobre o português afro-brasileiro de Lucchesi (2004, 2009d), em que foram encontrados, em relação à variável sexo, um leve favorecimento do sujeito nulo na fala dos homens, com 52 de peso relativo, enquanto há 48 de peso relativo na fala das mulheres, o que favorece levemente o uso do pronome sujeito. Nesse caso, encontramos as mulheres como responsável, provavelmente, pelo processo da mudança em curso.

As mulheres também são as que mais realizam o sujeito pronominal, e os homens os que mais utilizam o sujeito nulo também nos estudos sobre o português urbano culto de Duarte (1995), em que são encontradas 34% de ocorrências de sujeito nulo e peso relativo de 55 para o sexo masculino e 25% de ocorrências de sujeito nulo e 46 de peso relativo para o sexo feminino, o que vai se manter acentuado em seus trabalhos, embora, às vezes, se apresente menos acentuado em todos os grupos. O mesmo ocorreu em nossa pesquisa sobre o português popular brasileiro, em Novaes (2007, 2021), em que os resultados apontaram as mulheres com a tendência a uma maior utilização do sujeito pronominal pleno, com 59% de ocorrências do sujeito pronominal preenchido e peso relativo de 52, enquanto os homens apresentaram 43% de ocorrências e peso relativo de 47. Apesar da pequena diferença entre um fator e outro, esses resultados indicam que, provavelmente, as mulheres atuam de forma relevante no processo de mudança que vai na direção do favorecimento do uso do pronome sujeito na língua portuguesa no Brasil.

Sobre o papel das mulheres na mudança linguística, essa questão já foi ressaltada por Labov (2008 [1972]), ao afirmar que elas geralmente são responsáveis por encabeçar a mudança linguística. O mesmo resultado foi encontrado em uma pesquisa sobre o português paulista no século XIX, nos trabalhos de Costa (2020), em que há uma maior utilização, no geral, pelas mulheres, na realização do pronome de 1ª pessoa, dos que pelos homens, enquanto remetentes nas cartas que foram analisadas, sendo que há uma maior proporção de sujeitos nulos entre os homens. De acordo com Costa (2020), essa situação possivelmente indica que o uso do pronome pleno, que é considerado mais inovador no português brasileiro, já estava mais difundido entre as mulheres do que entre os homens.

Sendo assim, mais uma vez as mulheres aparecem liderando o processo de mudança. Isso, talvez, se deva ao fato de as mulheres estarem mais ligadas ao ambiente doméstico e à comunidade em que vivem, demonstrando certa reserva ao meio exterior. Por outro lado, os homens se deslocam para realizar trabalhos temporários ou mesmo em busca de empregos em outros centros urbanos, tendo assim uma maior mobilidade do que os falantes do sexo feminino. Eles mantêm contatos com outros grupos, se adaptando a novas situações, o que faz com que haja uma tendência à aquisição da norma linguística de outros grupos pelos falantes

desse sexo, principalmente quando se trata de comunidades rurais em que as mulheres da faixa etária III, mais de 65 anos, ainda mantêm um comportamento mais dedicado aos afazeres domésticos. Verificamos que as mulheres, estão à frente dos homens em relação ao preenchimento do pronome sujeito, nas três faixas etárias: 75,3% vs. 44,7%; na faixa etária I 63% vs. 59,1%; na faixa etária II e 62,4% vs. 39,8% na faixa etária III. Labov (2008 [1972]), no entanto, ressalta que dentro dessa questão, o mais importante não é se as mulheres ou se os homens lideram a mudança linguística, mas a observação de que a diferenciação entre os sexos irá desempenhar regularmente um importante papel no mecanismo da evolução linguística.

### 5.3.2 Faixa etária

Apesar de ter sido excluída na segunda rodada pelo Goldvarb, por não ter significância, apresentamos abaixo, na Tabela 9, a variável faixa etária. Verificamos na faixa etária II (de 45 a 55 anos) 60,6% de ocorrências de sujeito pronome, tendo em seguida a faixa I (de 25 a 35 anos) apresentado 56,8%, e a Faixa III (mais de 65 anos), 50,7%.

**Tabela 8** - A realização do sujeito pronominal expresso no português afro-brasileiro na comunidade quilombola de Lagoinha – Nova Canaã – BA, segundo a variável faixa etária.

Fatores	Aplicação/Total	Percentual
Faixa etária I	250/440	56,8
Faixa etária II	232/382	60,7
Faixa etária III	191/377	50,7
Total	673/1199	56,1

Fonte: Elaborada pelo autor

Essa observação é importante, porque percebemos um pequeno avanço das gerações mais jovens (faixas etárias I e II) para uma tendência ao preenchimento do

pronome sujeito e que esta variação já se encontrava nas gerações mais idosas (faixa etária III), pois o resultado percentual do uso ou não do pronome sujeito é praticamente o mesmo nos dois casos. Outro fato também considerável é que, apesar da pequena diferença entre as faixas etárias II e I, a segunda faixa etária poderia estar à frente, talvez pelo fato de terem mais contato com o mercado de trabalho e uma vida social, devido às responsabilidades assumidas em seu cotidiano.

### 5.3.3 Conclusão das variáveis sociais

Os resultados das variáveis sociais indicam que as mulheres estão à frente dos homens em se tratando da realização do pronome sujeito, apresentando peso relativo de 60, enquanto os homens apresentam 41. Esses resultados demonstram o favorecimento do preenchimento do pronome sujeito pelo fator sexo feminino o que provavelmente indica, mais uma vez, o papel das mulheres na propagação das variantes inovadoras e das mudanças linguísticas. Esse resultado nos mostra que, possivelmente, a mulher seria um dos agentes da mudança do parâmetro *pro-drop* no português popular do Brasil. Já em relação à faixa etária, apesar de não ter sido considerada significativa pelo programa computacional, demonstra um percentual que merece ser levado em consideração, que nos leva a considerar uma tendência a uma variação em relação ao preenchimento do sujeito, com as gerações mais jovens, II e I faixas etárias apresentando um percentual acima da geração da faixa etária III. Dessa forma, percebemos o avanço das gerações mais jovens (faixas etárias I e II), em relação ao preenchimento do pronome sujeito, sendo que essa variação já estava presente nas antigas gerações (faixa etária III)

## 5.4 CONCLUSÃO DA ANÁLISE DE DADOS

Os resultados da análise variacionista do preenchimento do sujeito pronominal no português popular afro-brasileiro na comunidade quilombola de Lagoinha revelaram uma maior utilização do pronome sujeito pleno em relação ao sujeito nulo. Dessa forma, verifica-se uma proximidade, em termos de ocorrências do pronome sujeito e do sujeito nulo, dos resultados encontrados nos trabalhos de Lucchesi (1994, 2009) sobre o português afro-brasileiro, Duarte (1995, 2019) sobre o

português urbano culto e Novaes (2007, 2021) sobre o português popular brasileiro. Por isso mesmo, apresentaram semelhanças no que diz respeito a alguns dos condicionantes linguísticos que regem a variação na fala de comunidades afro-brasileiras (LUCCHESI, 2004, 2009), na variedade urbana culta (DUARTE, 1995, 2019) e no português popular (NOVAES, 2007, 2021) do PB.

Nesses estudos, encontramos o favorecimento do pronome sujeito de acordo com a ordem de significância, segundo o programa Goldvarb X, através da variável menção anterior do sujeito, bem como sexo, pessoa do sujeito, caracterização semântica do sujeito e desinência gramatical do verbo, o que, de certo modo, se aproxima da afirmação de Lucchesi (2004, 2009) de que há o fortalecimento nas ações e nos processos de variação e mudança no sistema linguístico, o que ocorre tanto no plano diacrônico quanto sincrônico e nas suas variações diastráticas.

Em relação às ocorrências do pronome sujeito, em nossos estudos, mesmo que tenham sido realizadas em épocas e contextos diferentes, observamos que demonstraram resultados que se assemelham aos de Lucchesi (2004, 2009), Duarte (2005, 2019) e de nossa pesquisa anterior, Novaes (2007, 2021), no qual encontramos uma situação de favorecimento do preenchimento do pronome sujeito, e o que nos faz perceber uma situação de variação, que permite ao português brasileiro não mais se configurar como uma língua *pro-drop*.

Tanto em nossos estudos, quanto nos de Lucchesi (2004, 2009) sobre o português afro-brasileiro, assim como nos estudos sobre o português na norma culta (Duarte, 1995, 2019) como na norma popular, analisada nos estudos de Novaes (2007, 2021), percebe-se que o português brasileiro possivelmente está se afastando do grupo de línguas consideradas *pro-drop*, o que pode ser observado ao analisarmos os resultados da variável faixa etária, em que o processo de mudança é observado em tempo aparente, já que os mais jovens (faixa etária II e I) são os que mais realizam o sujeito pronominal pleno.

Há um estágio intermediário para o preenchimento do pronome sujeito na faixa etária III, sendo que há um leve favorecimento na faixa etária II, seguido pela faixa etária I, que se mantém dentro do nível da neutralidade. De qualquer forma, não deixa de ser um indicativo de que essa é uma tendência das faixas etárias mais jovens. Isso também ocorre tanto em Duarte (1995), quanto em Novaes (2007, 2021), em que os mais jovens (faixa etária I) são os que mais realizam o sujeito pronominal pleno. De acordo com Duarte (1995), o PB apresenta um sistema em

desenvolvimento em que a perda da “riqueza funcional” já não permite mais a identificação como uma língua com características *pro*, apesar de ainda conviver com um sistema que reflete características *pro-drop*. No entanto, outros estudos ainda precisam ser feitos para que esses resultados possam ser devidamente comprovados, já que muito ainda se tem a pesquisar sobre o português brasileiro.

## CONCLUSÕES

Esta pesquisa teve como objetivo principal fazer uma descrição da realização do sujeito pronominal na fala dos habitantes da comunidade quilombola de Lagoinha, situada no município de Nova Canaã, na Bahia. Para a elaboração deste estudo, foram consideradas 1.199 ocorrências de sujeito pronominal da fala de seis moradores de ambos os sexos, analfabetos e semianalfabetos, distribuídos em três faixas etárias: de 25 a 35 anos, de 45 a 55 anos e acima de 65 anos.

Cumprido ressaltar que a variabilidade do sujeito pronominal em uso na comunidade quilombola de Lagoinha foi investigada seguindo a orientação teórica da Sociolinguística, a qual atua nas fronteiras entre língua e sociedade, mostrando que há uma correlação entre fatores linguísticos e extralinguísticos que motivam e sistematizam o emprego das variantes.

Das variantes analisadas, verificamos a predominância do sujeito preenchido, considerando os resultados obtidos em pesos relativos para as 1.199 ocorrências de sujeito pronominal. Pode-se afirmar que o preenchimento do pronome sujeito tem como contextos favorecedores os seguintes fatores linguísticos e sociais em pesos relativos, segundo o grau de relevância: a primeira menção do sujeito com 66 de peso relativo; falantes do sexo feminino com 61 de peso relativo; a segunda pessoa do singular com peso relativo de 58; seguida pela primeira pessoa do singular, apresentando peso relativo de 54; o fator [+ animado / + humano], com peso relativo de 51; o morfema de 1ª pessoa com 61 de peso relativo.

O estudo da variação linguística, inegavelmente, enriqueceu-se com a utilização da metodologia da sociolinguística quantitativa laboviana. Hoje, sabemos que a variação linguística sofre efeitos de fatores internos ao sistema linguístico como externos (sexo, idade, estilo, dentre outros). Desse modo, podemos perceber a heterogeneidade da língua, sendo utilizada de maneiras diferentes entre os falantes.

A língua usada no Brasil, por exemplo, apesar de originária de Portugal, não pode ser considerada a mesma daquele país. São evidentes as diferenciações entre o que é falado no Brasil e o que é falado em Portugal, ou em qualquer outro país por ele colonizado. Ao chegar ao Brasil, o português misturou-se com a língua nativa, falada pelos indígenas e, mais tarde, com as dos negros escravizados vindos da

África. Desse contato, inúmeros processos variáveis surgiram influenciados pela questão temporal, espacial e social.

A língua portuguesa falada no Brasil perdeu algumas das características da variedade falada em Portugal, como a de língua *pro-drop*. A perda da morfologia flexional, ao comprometer, de certo modo, a estrutura funcional do português popular do Brasil, é um dos principais fatores nesse processo de mudança em curso. Nesse caso, no português do Brasil o preenchimento do sujeito é favorecido pela ausência de um morfema flexional, pelo mecanismo sintático da correferência e pelo traço semântico [+humano]. O mesmo pode ser observado, por exemplo, nos estudos de Duarte (1995, 2019), sobre o sujeito nulo no português culto, Lucchesi (2004), sobre o sujeito nulo no português afro-brasileiro, e Novaes (2007, 2021) sobre o português popular no interior do estado da Bahia.

Também podemos perceber que o contato entre línguas teve uma grande importância em relação à formação do português popular brasileiro. Não podemos negar a relevância do contato linguístico no período de colonização tanto no contexto histórico, quanto no contexto linguístico brasileiro. As evidências históricas demonstram que o PB foi influenciado por esse contato, principalmente o que se deu com as línguas africanas. Isso fica claro quando verificamos a influência dos meios de comunicação (rádio, TV e internet) e dos centros urbanos no português rural do Brasil atual. O estudo dos princípios e parâmetros e, mais especificamente, do parâmetro do sujeito nulo, vem nos mostrar o processo pelo qual a língua portuguesa no Brasil passou e vem passando. Apesar de nossos estudos não terem uma perspectiva paramétrica, discutimos neste trabalho o preenchimento do pronome sujeito, que é um dos parâmetros relacionados à Teoria de Princípios e Parâmetros. Nesse sentido, podemos verificar que o português do Brasil possui características próprias que o distinguem do PE.

Mesmo que haja pouco acesso à escolarização e à pequena influência que esses resultados poderiam ter, há possibilidade de certa interferência através de um contato maior dessas comunidades com o meio urbano, com o passar do tempo, o que também é complementado com a presença dos meios de comunicação (rádio, TV e, atualmente, a internet) na região. Isso ocorre na norma popular rural, de acordo com Lucchesi (2015), já que os mais jovens, através do contato com a cultura urbana, com o deslocamento dos moradores para buscarem opções de trabalho fora da comunidade, atividades comerciais etc., tendem a ter um contato

maior com a norma urbana culta, promovendo a substituição das antigas formas geradas pelo processo de contato linguístico, como a transmissão linguística irregular, por exemplo. Sendo assim, “Esse cenário de mudança observado nas comunidades rurais afro-brasileiras pode ser projetado, em sua essência, para todas as variedades populares do português brasileiro” (LUCCHESI, 2015, p.110).

Ao analisarmos os resultados da variável faixa etária, é possível perceber que os mais jovens (faixa etária II e I) são os que mais realizam o sujeito pronominal pleno, o que pode ser observado também tanto em Duarte (1995), quanto em Novaes (2007, 2021), em que os mais jovens (faixa etária I) são os que mais realizam o sujeito pronominal pleno. Isso nos leva a concluir que o português do Brasil está, cada vez mais, se afastando da propriedade *pro-drop*, que caracteriza línguas como o português europeu, o italiano e o espanhol. No entanto, outros estudos ainda precisam ser feitos para que esses resultados possam ser devidamente comprovados, já que muito ainda se tem a pesquisar sobre o PB.

Ao criar o modelo da teoria da variação e mudanças linguísticas, o objetivo de Labov foi o de responder como ocorre a mudança linguística. Desse modo, em relação ao português, é possível a formulação do seguinte princípio: se o português é tido como uma língua *pro-drop*, ou seja, de sujeito nulo, mas se há contextos em que o sujeito nulo não é favorecido, então, o processo de mudança da língua portuguesa ocorre justamente aí nesses contextos. E, em se tratando do português falado na comunidade quilombola de Lagoinha, pode-se depreender, diante dos resultados referentes aos pesos relativos, o encaixamento da mudança do sujeito nulo em direção ao preenchimento do pronome sujeito. Os resultados correspondem à nossa hipótese inicial de que haveria um maior preenchimento do pronome sujeito por se tratar de uma comunidade afrodescendente remanescente de quilombo, assim como, por outro lado, encontramos dados que indicam que o português brasileiro já não se configura mais como uma língua *pro-drop*.

É importante salientar que a análise descritiva do tema proposto não se esgota com este estudo. Seria desejável ampliar a amostra, a partir de uma estratificação mais detalhada, como também aprofundar a análise das variáveis linguísticas que se mostraram mais relevantes. O estudo desses aspectos, assim como a análise de outros fenômenos variáveis nessa mesma amostra, constitui possibilidades de desenvolvimentos futuros desta pesquisa, por isso mesmo sugerem o aprofundamento dessa discussão a respeito do preenchimento do

pronome sujeito, que tem se revelado um assunto de grande relevância para se compreender as mudanças ocorridas no português do Brasil, tendo como base a língua vernácula falada pelo povo brasileiro.

## REFERÊNCIAS

- ALMEIDA, Norma Lucia Fernandes de. *Sujeito nulo e morfologia verbal no português falado por três comunidades do interior da Bahia*. Tese de Doutorado. UNICAMP, Campinas-SP, 2005.
- ANJOS, Rafael S. A. *Territórios das comunidades remanescentes de antigos quilombos no Brasil: primeira configuração espacial*. 2. ed. Brasília: Mapas Editora e Consultoria, 2000.
- BALBI, Adriano. *Introduction à l'atlas ethnographique du globe*. Paris: Rey & Gravier, 1826.
- BARBOSA, Pilar; DUARTE, Maria Eugênia L.; KATO, Mary A. Null Subjects in European and Brazilian Portuguese. *Journal of Portuguese Linguistics*. Universidade de Lisboa, v. 4, p. 11-52, 2005.
- BAXTER, Alan N. A contribuição das comunidades afro-brasileiras isoladas para o debate sobre a crioulação prévia: um exemplo do estado da Bahia. In: d'ANDRADE, Ernesto; KIHM, Alain. *Actas do colóquio sobre crioulos de base lexical portuguesa*. Colibri, 1992, p. 7-35.
- BERWICK, Robert Cregar. *Locality Principles and the Acquisition of Syntactic Knowledge*. 451f. 1982. Dissertation (PhD) – Department of Electrical Engineering and Computer Science, MIT, Cambridge, 1982.
- BOECKX, C.A.; HORNSTEIN, N. On (Non-)Obligatory Control. In: DAVIES, W.D.; DUBINSKY, S. (eds.). *New Horizons in the Analysis of Control and Raising*. Studies in Natural Language and Linguistic Theory, v. 71. Springer, Dordrecht, 2007. [https://doi.org/10.1007/978-1-4020-6176-9\\_11](https://doi.org/10.1007/978-1-4020-6176-9_11)
- BONVINI, Emílio; PETTER, Margarida Maria Taddoni. Portugais du Brésil et langues africaines. *Langages*, v. 32, n. 130, p. 68-83, 1998.
- BONVINI, Emílio. Línguas africanas e português falado no Brasil. In: FIORIN, José Luiz; PETTER, Margarida. *África no Brasil: a formação da língua portuguesa*. São Paulo: Contexto. 2014, p. 15-62.
- BORGES NETO, José. *Ensaio de Filosofia da Linguística*. São Paulo: Parábola, 2004.
- BORTONI-RICARDO, Stella Maris. *Manual de Sociolinguística*. São Paulo: Contexto, 2014.
- BRIGHT, William. As dimensões da sociolinguística. In: FONSECA, M.S.V.; NEVES, M. F. (orgs.). *Sociolinguística*. Trad: E. N. Araújo Jorge. Rio de Janeiro: Eldorado, 1974 [1966], p. 17-24.

CALDEIRA, Jorge. História do Brasil com empreendedores. Juazeiro do Norte: Mameluco, 2009.

CALVET, Louis-Jean. *Sociolinguística: uma introdução crítica*. São Paulo: Parábola, 2002 [1992].

CARVALHO, Daniel da Silva; SOUSA, Lílian Teixeira de. Gramática Gerativa em perspectiva: escopo, objetivos e estrutura. In: CARVALHO, Daniel da Silva; SOUSA, Lílian Teixeira de (orgs.). *Gramática gerativa em perspectiva*. São Paulo: Blucher, 2018.

CHAMBERS, J. K.; TRUDGILL, Peter. *La Dialectología*. Trad. Carmen Morán González. Madrid: Visor Libros, 1994.

CHOMSKY, Noam. Topics in the theory of generative grammar. In: SEBEEK, T. (org.). *Currents Trends in Linguistics 3: linguistic theory*. Bloomington, Ind.: Indiana University Press, 1966.

CHOMSKY, Noam. *Aspectos da teoria da sintaxe*. Coimbra: Armênio Amado Editor, 1978 [1965].

CHOMSKY, Noam. *Some concepts and consequences of the theory of government and binding*. Chicago: MIT Press, 1982.

CHOMSKY, Noam. *Lectures on government and binding*. Dordrecht: Foris, 1988 [1981].

CHOMSKY, Noam. *O conhecimento da língua: sua natureza, origem e uso*. Lisboa: Editorial Caminho, 1994 [1986].

CHOMSKY, Noam. *Linguagem e mente: pensamentos atuais sobre antigos problemas*. Brasília: Editora Universidade de Brasília, 1998.

CHOMSKY, Noam. On the Nature, Use and Acquisition of Language. In: CLARK, Andy; TORIBIO, Josefa. (eds.). *Language and Meaning in Cognitive Science*. London: Routledge, 2012 [1987]. p. 13-32.

CHOMSKY, Noam. *Estruturas sintáticas*. Petropolis: Vozes, 2015 [1957].

CHOMSKY, Noam; LASNICK, Howard. Filters and control. *Linguistic Inquiry*, v. 8, p. 425-504, 1977.

COELHO, F. A. Os dialectos românicos ou neo-latinos na África. *Boletim da Sociedade de Geographia de Lisboa*. Lisboa, 1880. Reproduzido em: *Estudos linguísticos crioulos*. Lisboa, 1967, p. 129-196.

COELHO, F. A. Notas complementares. *Boletim da Sociedade de Geographia de Lisboa*. Lisboa, 1882. Reproduzido em: *Estudos linguísticos crioulos*. Lisboa, 1967, p. 451-478.

COELHO, F. A. Novas notas complementares. *Boletim da Sociedade de Geographia de Lisboa*. Lisboa, 1886. Reproduzido em: *Estudos linguísticos crioulos*. Lisboa, 1967, p. 705-755.

COELHO et al. *Para conhecer sociolinguística*. São Paulo: Contexto, 2015.

CYRINO, Sonia; DUARTE, Maria Eugênia L.; KATO, Mary. Visible subjects and invisible clitics in Brazilian Portuguese. In: KATO, M. A.; NEGRÃO, E. V. (orgs.). *Brazilian Portuguese and the Null Subject Parameter*. Madrid: Iberoamericana/Frankfurt: Vervuert, p. 55-73, 2000.

DUARTE, Maria Eugênia Lamoglia. Clítico acusativo, pronome lexical e categoria vazia no português do Brasil. In: TARALLO, Fernando (org.). *Fotografias sociolinguísticas*. Campinas: Pontes. 1989, p. 19-34.

DUARTE, Maria Eugênia Lamoglia Do pronome nulo ao pronome pleno: a trajetória do sujeito no português do Brasil. In: ROBERTS, Ian; KATO, Mary A. (org.). *Português brasileiro: uma viagem diacrônica*. Campinas: Pontes, 1993, p. 107-128.

DUARTE, Maria Eugênia Lamoglia. *A perda do princípio “evite pronome” no português brasileiro*. 1995. 151f. Tese (Doutorado em Linguística) – Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, Campinas, SP, 1995.

DUARTE, Maria Eugênia L.; MOURÃO, Gabriela C.; SANTOS, Heitor M. Os sujeitos de 3ª pessoa; revisitando Duarte. In: DUARTE, Maria Eugênia Lamoglia (org.). *O sujeito em peças de teatro (1833-1992): estudos Diacrônicos*. São Paulo: Parábola, 2012, p. 21-44.

DUARTE, Maria Eugênia. Sociolinguística “Paramétrica”. In: MOLLICA, Maria Cecília; FERRAREZI Jr., Celso (orgs). *Sociolinguística, sociolinguísticas: uma introdução*. São Paulo: Contexto, 2016.

DUARTE, Maria Eugênia Lamoglia. O sujeito nulo referencial no português brasileiro e no português europeu. In: GALVES, Charlotte; KATO, Mary A.; ROBERTS, Ian (org.). *Português brasileiro: uma segunda viagem diacrônica*. Campinas: Editora da Unicamp, 2019, p. 93-126.

EBERHARD, David M.; SIMONS, Gary F.; FENNIG, Charles D. *Ethnologue: Languages of the World*. 25th edn. Dallas: SIL International. 2022.

ELIA, Sílvio. *A unidade linguística do Brasil*. Rio de Janeiro: Padrão, 1979.

GALVES, Charlotte A sintaxe do português brasileiro. *Ensaio de Linguística*, v. 13, p. 31-50, 1987.

GALVES, Charlotte A sintaxe pronominal do português brasileiro e a tipologia dos pronomes. *Ensaio sobre as gramáticas do português*. Campinas: Editora da Unicamp, 2001 [1987], p. 152-179.

Galves, C. *Ensaio sobre as gramáticas do Português*. Campinas: Editora da UNICAMP, p. 43-58, 2001.

GALVES, Charlotte. Revisitando a concordância no português brasileiro. In: GALVES, C.; KATO, M. A.; ROBERTS, I. *Português brasileiro: uma segunda viagem diacrônica*. Campinas: Unicamp, 2019, p. 127-150.

GRÉDIS, Ana Rosi. A importância dos estudos sobre a Gramática Universal nas pesquisas em aquisição de segunda língua. *Estudos Linguísticos e Literários*, n. 44, p. 163-181, 2016.

GUY, Gregory Riordan. *Linguistic variation in Brazilian Portuguese: aspects of phonology, syntax and language history*. PhD Dissertatiion – University of Pennsylvania, Philadelphia, 1981.

GUY, Gregory Riordan. On the nature and origins of popular Brazilian Portuguese. *Estudios sobre Español de America y Lingüística Afroamericana*. Bogotá: Instituto Caro y Cuervo, p. 227-245, 1989.

GUY, Gregory Riordan; ZILES, Ana Marial. *Sociolinguística quantitativa: instrumental de análise*. São Paulo: Parábola, 2007 [1993].

HOLM, John. Creole influence on popular Brazilian Portuguese. In: GILBERT, G. (ed.). *Pidgin and Creole Language*. Honolulu: University of Hawaii Press, 1987, p. 406-429.

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

HUANG, James. On the distribution and reference of empty pronouns. *Linguistic Inquiry*, v. 15, n. 4, p. 531-574, 1984.

HYAMS, Nina. The Pro-drop parameter in child grammars. In: *Proceedings of the West Coast Conference on Formal Linguistics*, v. 2, 1983.

IBGE. *Nova Canaã - História e fotos*. Disponível em: <<https://cidades.ibge.gov.br/brasil/ba/nova-canaa/historico>>. Acesso em 21 de dezembro de 2021.

IBGE. *Nova Canaã - Panorama*. Disponível em: <<https://cidades.ibge.gov.br/brasil/ba/nova-canaa/panorama>>. Acesso em 21 de dezembro de 2021.

JERNE, Niels. K. Antibodies and learning: selection versus instruction. In: QUARTON, G. C.; MELNECHUK, T.; SCHMITT, F. O. (eds.). *The Neurosciences: a study program*. Rockefeller University Press, 1967.

JESPERSEN, Otto. *The philosophy of grammar*. New York: Norton, 1935 [1924].

KATO, M. Strong and weak pronominals in the null subject parameter. *Probus*, vol. Nº11, p. 1-37, 1999.

KATO, Mary A. A evolução da noção de Parâmetros. *DELTA*, São Paulo, v. 18. n. 2, p. 309-337, 2002.

KATO, Mary. Aspectos morfofonológicos nos paradigmas dos pronomes fortes e fracos. *Revista da Anpoll*, n. 46, v. 1, p. 142-155, 2018.

KATO, Mary. O português brasileiro e o sistema de referência nas línguas naturais. *Estudos linguísticos e literários*, n. 72, p. 265-270, 2021.

KROCH, A. Reflexes of grammar in patterns of language change. *Language Variation and Change*. v.1, p.199–244, 1989.

KROCH, Anthony. Syntactic Change. In: BALTIN, Mark; COLLINS, Chris (orgs.) *The Handbook of Contemporary Syntactic Theory*. Oxford: Blackwell, 2001, p. 699-729.

KUHN, Thomas. Reflections on my critics. In: LAKATOS, I.; MUSGRAVE, A. (eds.). *Criticism and the growth of knowledge: proceedings of the International Colloquium in the Philosophy of Science*. London: Cambridge University Press, 1970.

LABOV, William. The Linguistic Variable as a Structural Unit. *Washington Linguistic Review*. v. 3, 1966, p. 1-22.

LABOV, William *Padrões sociolinguísticos*. Tradução: Marcos Bagno, Maria Marta Pereira Scherre, Caroline Rodrigues. São Paulo: Parábola Editorial, 2008 [1972].

LABOV, William. On the use of the present to explain the past. In: KIEFERE, Ferenc; VAN STERKENBURG, Piet (eds.). *Eight Decades of General Linguistics*. Leiden: Brill, 2013 [1975].

LIGHTFOOT, David. The child's trigger experience: degree-0 learnability. *Behavioral and Brain Sciences*, v. 12, 2. ed., p. 321-334, 1989.

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

LOBATO, Lúcia Maria Pinheiro. A teoria da regência e ligação. In: LOBATO, Lúcia Maria Pinheiro. *Sintaxe gerativa do português: da teoria padrão à teoria da regência e ligação*. Belo Horizonte: Vigília, capítulo XII, 1986, p. 399-495.

LUCCHESI, Dante. O conceito de transmissão linguística irregular e o processo de formação do português do Brasil. In: RONCARATI, Cláudia; ABRAÇADO, Jussara (orgs.). *Português brasileiro: contato linguístico, heterogeneidade e história*. Rio de Janeiro: 7 Letras, 2003, p. 272-284.

LUCCHESI, Dante. Contato entre línguas e variação paramétrica: o sujeito nulo no português afro-brasileiro. *Língua(gem)*, Macapá, v. 1, n. 2, 2004, p. 63-91.

LUCCHESI, Dante. Introdução. In: LUCCHESI, Dante; BAXTER, Alan; RIBEIRO, Ilza (orgs.). *O português afro-brasileiro*. Salvador: Edufba, Introdução, 2009a, p. 27-37.

LUCCHESI, Dante. História do contato entre língua no Brasil. In: LUCCHESI, Dante; BAXTER, Alan; RIBEIRO, Ilza (orgs.). *O português afro-brasileiro*. Salvador: Edufba, cap. 1, 2009b, p. 41-73.

LUCCHESI, Dante. Metodologia. In: LUCCHESI, Dante; BAXTER, Alan; RIBEIRO, Ilza (orgs.). *O português afro-brasileiro*. Salvador: Edufba, 2009c, p. 155-164.

LUCCHESI, Dante. A realização do sujeito pronominal. In: LUCCHESI, Dante; BAXTER, Alan; RIBEIRO, Ilza (orgs.). *O português afro-brasileiro*. Salvador: Edufba, cap. 6, 2009d, p. 167-183.

LUCCHESI, Dante. Os limites da variação e da invariância na estrutura da gramática. *Revista da ABRALIN*, v. Eletrônico, n. Especial, p. 227-259, 2ª parte, 2011.

LUCCHESI, Dante. *Língua e sociedade partidas: a polarização sociolinguística do Brasil*. São Paulo: Contexto, 2015.

MACEDO SOARES, A. J. de. Estudos lexicográficos do dialecto brasileiro IV: Sobre algumas palavras africanas introduzidas no portuguez que se fala hoje no Brazil. *Revista Brasileira*, tomo IV, 2, p. 243-271, 1880.

MACHADO, Elenir de Oliveira (2002). Morfema zero e casos de alomorfia. In: SILVA, José Pereira da (org.). Morfossintaxe da língua portuguesa. *Cadernos da pós-graduação em Língua Portuguesa*, n. 2, Suplemento. São Gonçalo: Faculdade de Formação de Professores/UERJ, p. 7-13.

MANESSY, Gabriel. *Créoles, pidgins, variétés véhiculaires, procès et genèse*. Paris: Editions du CNRS, 1995.

MATOS, Noeme Martins; OLIVEIRA, Agnaldo Menezes de. Coletâneas do Colégio Florestal. Nova Canaã: [S.I.], 1999.

MATTOS E SILVA, Rosa Virgínia. *Ensaio para uma sócio-história do português brasileiro*. São Paulo: Parábola, 2004.

MELO, Gladstone Chaves de. *A língua do Brasil*. Rio de Janeiro: Padrão, 1981 [1946].

MENDONÇA, Renato. *Influência Africana no Português do Brasil*. Brasília: Fundação Alexandre Gusmão, 2012 [1933].

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

MOLLICA, Maria Cecília; BRAGA, Maria Luiza (orgs.). *Introdução à Sociolinguística: o tratamento da variação*. 2. ed. São Paulo: Contexto, 2004.

MORALES, Humberto López. *Sociolingüística*. Madrid: Editorial Gredos, 2015 [1993].

MUSSA, Alberto B. N. *O papel das línguas africanas na história do português do Brasil*. 1991. Dissertação (Mestrado em Linguística) – Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 1991.

NIYOGI, Partha; BERWICK, Robert. Evolutionary consequences of language learning. *Linguistics and Philosophy*, n. 20, p. 697-719, 1997.

NOVAES, José Carlos Assunção. *O parâmetro do sujeito nulo no português popular do interior do estado da Bahia*. 2007. 190f. Dissertação (Mestrado em Letras) – Instituto de Letras, Universidade Federal da Bahia, 2007.

NOVAES, José Carlos Assunção. *O parâmetro do sujeito nulo no português popular da Bahia*. São Paulo: Paco Editorial, 2021.

PAGOTTO, Emílio. Ebulição e sedentarização linguística: o lugar da economia de subsistência na formação do português brasileiro. *Diadorim*, v. 20, p. 53–63, 2018.

PAIVA, Maria da Conceição. A variável gênero/sexo. In: MOLLICA, M. Cecília; BRAGA, M. L. (orgs.). *Introdução à Sociolinguística: o tratamento da variação*. 2. ed. São Paulo: Contexto, 2004, p. 33-42.

PESSOA DE CASTRO, Yeda. *De l'intégration des apports africains dans les parles de Bahia au Brésil*. Tese (Doutorado em Letras) – Université Nationale du Zaïre, 2 v. 1976.

PESSOA DE CASTRO, Yeda. *Os falares africanos na interação social do Brasil Colônia*. Salvador, Centro de Estudos Baianos/UFBA, n. 89. 1980.

PESSOA DE CASTRO, Yeda. Os falares africanos na interação social dos primeiros séculos. In: MELLO, Linalda Arruda (Org.). *Sociedade, cultura & língua*. João Pessoa: Shorin, 1990, p. 91-113.

PESSOA DE CASTRO, Yeda. *Falares africanos na Bahia*. 2. ed. Rio de Janeiro: Topbooks, 2005.

PONTE, Vanessa. *A indeterminação do sujeito no português popular do interior do estado da Bahia*. 2008. 127f. Dissertação (Mestrado em Letras) – Instituto de Letras, Universidade Federal da Bahia, Salvador, 2008.

PONTES, Eunice. Da importância do tópico em português. *Anais do V Encontro de Lingüística*, v. II. Rio de Janeiro: PUC, 1981.

QUINE, W.V. Reflexões metodológicas sobre a teoria lingüística atual. *Síntese* 21 (3-4), 1970, p. 386-398.

RANGLES, Willaim Graham Lister. *L'ancien royaume du Congo des origines à la fin du XIX<sup>e</sup> siècle*. Paris/La Haye: Mountain, 1968.

- RAPOSO, Eduardo Paiva. *Teoria da Gramática: a Faculdade da Linguagem*. Lisboa: Caminho, 1992.
- RAYMUNDO, Jacques. *O elemento afro-negro na língua portuguesa*. Rio de Janeiro: Renascença. 1933.
- RIBEIRO, João. *Diccionario grammatical*, 3. ed. inteiramente refundida e muito augmentada. Rio de Janeiro/ São Paulo/ Belo Horizonte: Francisco Alves. 1906 [1888].
- RIZZI, L. *Issues in Italian syntax*. Dordrecht: Foris, 1982.
- ROBERTS, Ian. Posfácio. O português brasileiro no contexto das línguas românicas. In: ROBERTS, I.; KATO, M. A. (orgs.). *Português brasileiro: uma viagem diacrônica*. Tradução: Ruth Lopes. Campinas, SP: Ed. da Unicamp, 1993, p. 401-421.
- ROBERTS, Ian. Creoles, markedness and the Language Bioprogram Hypothesis. *Estudos Linguísticos e Literários*, Salvador, n. 19, p. 11-24, 1997.
- ROBERTS, Ian. *Diachronic Syntax*. New York: Oxford University Press, 2007.
- RODRIGUES, R.N. *Os africanos no Brasil*. São Paulo: Companhia Editora Nacional. 1935.
- SAUSSURE, Ferdinand de. *Curso de Linguística Geral*. São Paulo: Cultrix, 22. ed, 2000 [1916].
- SCHWARTZ, Stuart B. Mocambos, Quilombos e Palmares: A resistência escrava no Brasil colonial. *Estudos Econômicos*. São Paulo, v. 17, n. 9, Especial, 1987, p. 61-88.
- SCHWARTZ, Stuart B. *Cantos e quilombos numa conspiração de escravos haussás*. Bahia, 1814. In: Reis, JJ; Gomes, F. S. (orgs.). *Liberdade por um fio: história de quilombos no Brasil*. São Paulo: Companhia das Letras, 1994, p. 373-400.
- SILVA, Jorge Augusto Alves da. (2005) *A concordância verbal de terceira pessoa do plural no português popular do Brasil: um panorama sociolingüístico de três comunidades do interior do estado da Bahia*. Salvador: UFBA. Tese de doutorado.
- SILVA, Rosemary Francisca Neves. Mulheres negras no Brasil Colonial e a privação da vivência de sua intimidade. (org) LEMOS, Carolina Teles. In: *Religião e (Re)significação da Intimidade*. Goiânia: Ed. Da PUC Goiás; Ed. Kelps, 2012, p. 89-100.
- SILVA-CORVALÁN, Carmem. *Sociolingüística, teoría y análisis*. Madrid: Alhambra, 1988.
- SILVA NETO, Serafim da. *Introdução ao estudo da língua portuguesa no Brasil*. 2. ed. Rio de Janeiro: INL, 1963 [1951].

TARALLO, Fernando. *Relativation strategies in Brazilian Portuguese*. 1983. 273f. PhD Dissertation – University of Pennsylvania, Philadelphia, 1983.

TARALLO, Fernando. Por uma sociolingüística românica ‘paramétrica’: fonologia e sintaxe. *Ensaio de Lingüística*. n. 13, p. 51-83, 1987.

TARALLO, F. Reflexões sobre o conceito de mudança lingüística. *Organon*, n. 18, p. 11-22, 1991.

TARALLO, Fernando; KATO, Mary. Harmonia Trans-sistêmica: variação intra e inter-lingüística. *Revista Diadorim*. Rio de Janeiro: Ed. UFRJ, v. 2, p. 13-42, 2007 [1989].

TARALLO, Fernando. Diagnosticando uma gramática brasileira: o português d’aquém e d’além mar ao final do século XIX. In: ROBERTS, Ian; KATO, Mary A. (orgs.). *Português brasileiro: uma viagem diacrônica*. 3.ed. São Paulo: Contexto, 2018a [1993], p. 55-82.

TARALLO, Fernando. Sobre a alegada origem crioula do português brasileiro: mudanças sintáticas aleatórias. In: ROBERTS, Ian; KATO, Mary A. (Orgs.). *Português brasileiro: uma viagem diacrônica*. 3.ed. São Paulo: Contexto, p. 29-53, 2018b [1993].

TAUNAY, Afonso d’E. *Subsídio para a História do Tráfico Africano no Brasil*. São Paulo: Imesp, 1941.

VIANA FILHO, Luiz. *O negro na Bahia*. Prefácio de G. Freyre. Rio de Janeiro: José Olympio, 1946.

WEINREICH, Uriel; LABOV, William; HERZOG, Marvin I. *Fundamentos empíricos para uma teoria da mudança lingüística*. Tradução: Marcos Bagno. São Paulo: Parábola Editorial, 2006 [1968].

XAVIER, Gildete Rocha. *Português brasileiro como segunda língua: um estudo sobre o sujeito nulo*. 2006. 213f. Tese (Doutorado em linguística) – Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, Campinas, SP, 2006.