



O português do Oeste baiano: constituição de *corpus* em Santa Maria da Vitória

The Portuguese of Western Bahia: corpus constitution in Santa Maria da Vitoria

Jéssica Carneiro da Silva*
Universidade do Estado da Bahia
Universidade Federal da Bahia
Salvador, Bahia, Brasil

Maria Cristina Vieira de Figueiredo Silva**
Universidade Federal da Bahia
Salvador, Bahia, Brasil

Gilianderson Castro da Silva***
Universidade Federal do Oeste da Bahia
Santa Maria da Vitória, Bahia, Brasil

Thalita Oliveira Fernandes de Araújo****
Universidade Federal do Oeste da Bahia
Santa Maria da Vitória, Bahia, Brasil

Ísis Juliana Figueiredo de Barros****
Universidade Federal do Recôncavo da Bahia
Amargosa, Bahia, Brasil

Resumo: Este trabalho tem como objetivo discutir aspectos metodológicos de constituição de *corpus* e, para tanto, apresenta a metodologia realizada pelo projeto “Os falares do Além São Francisco” (FIGUEIREDO, 2015), da UFBA, em parceria com o projeto “O português rural do Oeste da Bahia”, da UFOB (BARROS, 2015), bem como o levantamento de dados linguísticos utilizados como objeto de pesquisa da tese “O português falado no Oeste baiano: constituição de corpus e análise das estratégias de relativização. O objetivo central da pesquisa é o de contribuir para a compreensão e descrição linguísticas dos falares baianos do Oeste da Bahia e da sua formação sócio-histórica, através da constituição de *corpus* em Santa Maria da Vitória (SAMAVI). A partir da coleta da fala vernacular da comunidade quilombola de Montevidinha e da comunidade

* Doutoranda em Língua e Cultura pela Universidade Federal da Bahia, professora substituta da Universidade do Estado da Bahia, *campus* Alagoinhas. E-mail: jessxcs@gmail.com ou jecsilva@uneb.br.

** Professora do Programa de Pós-Graduação em Língua e Cultura da Universidade Federal da Bahia. E-mail: macrisfig@uol.com.br.

*** Graduando em Publicidade e Propaganda pela Universidade Federal do Oeste da Bahia. E-mail: gilicastro2017@gmail.com.

**** Graduanda em Artes Visuais pela Universidade Federal do Oeste da Bahia. E-mail: thalitafern@gmail.com.

**** Professora Doutora do Centro de Formação de Professores da Universidade Federal do Recôncavo da Bahia. E-mail: isis.barros@ufrb.edu.br.

urbana de SAMAVI, ambas localizadas no Oeste baiano, desenvolve-se a análise de estudos sociolinguísticos e sócio-históricos, resgatando a importância da região, tendo em vista os dados históricos percebidos por meio das narrativas orais, linguísticas, culturais e sociais, que serão apresentados brevemente neste trabalho. Para a construção metodológica da pesquisa, toma-se como base a Sociolinguística Laboviana, por meio do levantamento de células sociais, entrevistas e transcrições linguísticas. Até o momento, nossos resultados abrangem: 12 entrevistas (11 concluídas e uma em andamento) dos falantes de Montevidinha; em SAMAVI, foram gravadas 12 entrevistas sociolinguísticas, reunidas em um *corpus* com a seguinte estratificação social: três faixas etárias (Faixa I – 25 a 35 anos, Faixa II – 45 a 55 anos, Faixa III – acima de 65 anos) e dois sexos (masculino e feminino). Espera-se que o estudo contribua para a compreensão e descrição linguísticas dos falares do Oeste da Bahia, além de contribuir para a documentação de sua formação sócio-histórica, para a ampliação dos estudos sociolinguísticos no Brasil e para a compreensão da formação sócio-histórica do português brasileiro.

Palavras-chave: Sociolinguística. Constituição de *corpus*. Santa Maria da Vitória. Português do Oeste baiano.

Abstract: This work aims to discuss methodological aspects of the constitution of the corpus and, therefore, presents the methodology carried out by the project "The speeches of the beyond São Francisco" (FIGUEIREDO, 2015), from UFBA, in partnership with the project "The rural Portuguese of the Oeste da Bahia", by UFOB (BARROS, 2015), as well as the survey of linguistic data used as the research object of the thesis "Portuguese spoken in western Bahia: constitution of corpus and analysis of relativization strategies. This research expects to contribute to the understanding and linguistic description of speakers from the Western Bahia and their socio-historical formation, through the constitution of *corpus* in Santa Maria da Vitória (SAMAVI) and Montevidinha. With the collecting of the vernacular speech of the quilombola community of Montevidinha and SAMAVI, the research aims the analysis of sociolinguistic and sociohistorical studies, rescuing the importance of the region, having in view of historical data perceived through oral, linguistic, cultural and social narratives. For the methodological construction of the research, we based on the Labovian Sociolinguistics approach, through the lifting of social cells, interviews and linguistic transcriptions. Our results cover 12 interviews (11 completed and 1 in progress) of the speakers of Montevidinha. In SAMAVI, 12 completed sociolinguistic interviews were recorded in a *corpus* with the following social stratification: three age groups (Group I – 25-35 years old, Group II – 45-55 years old, and Group III – above 65 years old) and two sexes (men and women). The study can contribute to the understanding and linguistic description of the speeches of the West of Bahia. It also contributes to the documentation of their socio-historical formation, to the expansion of sociolinguistic studies in Brazil and the understanding of the socio-historical formation of the Brazilian Portuguese.

Keywords: Sociolinguistics. *Corpus* constitution. Santa Maria da Vitória. Portuguese of Western Bahia.

1 INTRODUÇÃO

As mudanças linguísticas ocorridas no português brasileiro (PB) podem ser explicadas pelo intenso contato linguístico entre falantes do português e falantes de línguas indígenas e africanas durante os períodos colonial e imperial (MATTOS E SILVA, 2004; LUCCHESI, 2009), o que justifica a realização de pesquisas em comunidades em que se observa em sua história a presença de indígenas e afrodescendentes. Tem-se adotado que as mudanças ocorridas no PB se devem a um processo de transmissão linguística irregular (TLI), hipótese proposta inicialmente por Baxter (1995) e reiterada por Lucchesi e Baxter (2009). A TLI configura-se como "um contínuo de níveis diferenciados de socialização/nativização de uma língua segunda, adquirida massivamente, de forma mais ou menos imperfeita, em contextos sócio-históricos específicos" (LUCCHESI, 2003, p. 104).

Assume-se a posição de que o contexto multilinguístico existente na época da colonização foi o desencadeador das estratégias inovadoras do PB, que o fez se distanciar do português europeu (PE), devido ao precário contato entre falantes de línguas distintas com a língua alvo. Parte-se da hipótese de que essa simplificação tenha ocorrido também em Santa Maria da Vitória (SAMAVI) e Montevidinha, comunidades de fala desta pesquisa que estão localizadas no Oeste da Bahia, já que há indícios e evidências da presença indígena e africana nessa região. Fica evidente, então, “a importância de se considerar a transmissão linguística irregular na formação da identidade sociolinguística brasileira, aplicando-a não apenas no tocante à participação africana mas também à indígena” (ARAÚJO, 2017, p. 50).

Além do fato de as comunidades possuírem características que subsidiam a discussão a respeito da formação do PB, houve a constatação de que não há amostras de fala do Norte nem do Oeste baiano ou de pesquisas que contemplem registros sociolinguísticos dessas regiões. Existem projetos que contemplam outras regiões da Bahia, tais como: i. o projeto *Programa de Estudos do Português Popular Falado de Salvador* (PEPP), coordenado pelo professora Norma da Silva Lopes, que reuniu amostras de fala do português popular de Salvador entre 1998 e 1999; ii. o projeto *Vertentes do português popular do estado da Bahia*, coordenado pelo Professor Dante Lucchesi (UFBA), que trabalha registrando e analisando a fala de comunidades rurais afro-brasileiras relativamente isoladas distribuídas na Chapada Diamantina (Barra, Bananal, na região de Rio de Contas), no Sul do estado (Helvécia), no Semiárido (Cinzento) e no Recôncavo baiano (Sapé), de comunidades rurais não marcadas etnicamente (Santo Antônio de Jesus e Poções) e de comunidades urbanas da região metropolitana de Salvador (Liberdade, Itapuã, Plataforma, Cajazeiras e Lauro de Freitas); iii. o projeto *A língua portuguesa no semiárido baiano*, sob a coordenação das professoras Norma Lúcia Fernandes de Almeida, Eliana Pitombo Teixeira, Zenaide Oliveira Novaes Carneiro e Silvana Silva de Farias de Araújo (UEFS), que estuda as comunidades de Piemonte da Diamantina, Rio de Contas, Jeremoabo (Chapada Diamantina) e Paraguaçu (Feira de Santana), além de amostras das normas culta e popular do português falado em Feira de Santana.

Assim, a relevância deste artigo¹ justifica-se por: i. contribuir para os estudos sociolinguísticos no Brasil, ao suscitar reflexões sobre a formação sócio-histórica do PB, descrevendo e analisando sua diversidade sociolinguística; ii. cooperar na constituição de corpus de SAMAVI, a fim de posteriormente descrever e explicar a variação e a mudança nas estratégias de relativização, um fenômeno que distingue o PB do PE.

A tese de Carneiro (no prelo) tem como objetivos gerais contribuir para a compreensão e descrição linguísticas dos falares baianos do Oeste da Bahia e da sua formação sócio-histórica, através da constituição de *corpus* em Santa Maria da Vitória e investigar o uso variável das estratégias de relativização, analisando todas as relativas desenvolvidas tanto nas amostras de fala de Santa Maria da Vitória (SAMAVI) quanto nas

¹ Este artigo apresenta resultados parciais da tese de doutorado em andamento da primeira autora, orientada pela segunda autora, intitulada “O português falado no Oeste baiano: constituição de *corpus* e análise das estratégias de relativização.

de Montevidinha². E tem os seguintes objetivos específicos: i. constituir *corpus* com amostras de fala de Santa Maria da Vitória (sede) e de Montevidinha (zona rural), ambas localizadas no Oeste baiano; ii. realizar análise das estratégias de relativização nas duas comunidades mencionadas, seguindo o aporte teórico e metodológico da Sociolinguística Variacionista; iii. compreender quais fatores linguísticos e/ou sociais condicionam e/ou favorecem as relativas no português Oeste baiano; iv. mapear o uso das relativas no Estado da Bahia, descrevendo e confrontando com trabalhos realizados sobre o tema nos diversos falares e localidades do estado baiano; v. caracterizar o português falado no Oeste baiano, contribuindo, assim, para o conhecimento da realidade sociolinguística do PB; vi. fornecer subsídios a favor da hipótese da transmissão linguística irregular para explicar a formação do PB como resultado do contato linguístico; vii. documentar dados bibliográficos e demográficos sobre a sócio-história da região, produzindo um acervo documental; viii. reunir depoimentos dos informantes das comunidades e dos registros acessíveis em órgãos públicos.

Para este artigo, objetiva-se principalmente discutir aspectos metodológicos de constituição de corpus e, para tanto, são apresentados a metodologia desenvolvida nos projetos e o levantamento de dados linguísticos utilizados como objeto de pesquisa da tese. Além de apresentar a metodologia utilizada no desenvolvimento dos inquéritos, este trabalho objetiva descrever a sócio-história do Oeste da Bahia, focando especificamente na cidade de Santa Maria da Vitória e na comunidade rural de Montevidinha, comunidades de fala em que os inquéritos foram realizados.

Na primeira seção, desenvolve-se uma discussão em torno da literatura sociolinguística, pontuando noções da teoria da variação e da mudança linguística necessárias para o desenvolvimento da metodologia de uma pesquisa variacionista e quantitativa. Nessa mesma seção, discute-se brevemente a metodologia laboviana quantitativa usada nos estudos variacionistas. A segunda seção descreve a região do Oeste da Bahia, as comunidades de fala (SAMAVI e Montevidinha) e sua sócio-história, a partir de uma pesquisa bibliográfica realizada durante a ida a campo para a constituição do *corpus* e o levantamento dos dados. Na terceira seção, são descritos os aspectos metodológicos da constituição de *corpus* em SAMAVI e do desenvolvimento dos inquéritos já realizados, abordando aspectos como: levantamento e seleção de informantes, quadro de informantes entrevistados, estratificação social considerada (as variáveis sociais), procedimentos éticos, tipo de entrevista, recursos tecnológicos utilizados e possíveis variáveis linguísticas para análise futura. Por fim, apresentam-se os resultados esperados com a pesquisa e a partir da constituição de *corpus*, que podem elucidar contribuições tanto para a pesquisa sociolinguística brasileira quanto para a ampliação da documentação sócio-histórica sobre a região.

² O objetivo da referida tese em andamento (CARNEIRO, no prelo) é o de colaborar na constituição de corpus apenas em SAMAVI. O corpus da comunidade rural de Montevidinha já foi anteriormente constituído por pesquisadores do projeto “O português rural do Oeste da Bahia”, vinculado à UFOB, coordenado pela Profa. Dra. Ísis Juliana Figueiredo de Barros, à época, professora adjunta na instituição. Agradecemos também as contribuições valiosas do museólogo Hermes Novais, e colaboração dos parceiros Levi Rodrigues, Prof. Me. Marcelo da Silva Souza e Profa. Dra. Manan Terra Cabo.

Para além da descrição e análise da língua (possíveis com o desenvolvimento de um *corpus* de dados reais de fala), o labor sociolinguístico, em essência, fundamenta-se na afirmação de que a fala dos informantes é moldada – e por isso deve ser também explicada – pelo contexto sócio-histórico e geográfico, no entendimento de que “o falante encerra em si a síntese do geográfico com o social” (CARDOSO, 2010, p. 12). Portanto é na fala que se encontram as marcas identitárias, sociais, históricas e culturais de um povo.

2 A TEORIA SOCIOLINGÜÍSTICA E A SUA METODOLOGIA QUANTITATIVA

Com vistas a descrever a metodologia da pesquisa sociolinguística desenvolvida no Oeste da Bahia, em SAMAVI e Montevidinha, faz-se necessário discorrer sobre a teoria que fundamenta os aspectos metodológicos tratados neste artigo. Dessa forma, explicita-se a teoria da variação e da mudança linguística, pontuando conceitos necessários para o entendimento e o direcionamento metodológico de toda e qualquer pesquisa quantitativa laboviana.

2.1 A TEORIA DA VARIAÇÃO E DA MUDANÇA LINGÜÍSTICA

A teoria da variação e da mudança linguística, postulada por Weinreich, Labov e Herzog (2006 [1968]), é o ramo da Linguística que estuda as variações e as mudanças linguísticas dentro de uma comunidade de fala. Tem por objeto de estudo os padrões observáveis do comportamento linguístico dentro de uma comunidade de fala e os formaliza analiticamente através de um sistema heterogêneo, constituído por unidades e regras variáveis (WEINREICH; LABOV; HERZOG, 2006 [1968]). Para essa teoria, a mudança deve ser compreendida considerando a vida social da comunidade em que são produzidos os dados de fala, pois a língua sofre pressões e influências sociais constantemente. Postula-se, pois, a impossibilidade de perceber a mudança linguística fora da comunidade de fala e admite-se que ela não é proveniente de uma deriva natural, mas sim fruto da variação inerente na comunidade de fala, motivada pelo encaixamento linguístico e social das variantes linguísticas.

Além de propor a contextualização da mudança linguística na sociedade, a Sociolinguística Variacionista propõe o postulado de que, para analisar a variação e/ou mudança linguística, é fundamental considerar explicações linguísticas e extralinguísticas. Ou seja, é preciso que se investiguem os fatores que condicionam a escolha do falante pelas variantes em competição a partir da observação dos padrões coletivos de comportamentos linguísticos numa dada comunidade de fala, a qual é interpelada por fatores linguísticos e sociais.

De acordo com Labov (1994 [1972], p. 19), “a mudança linguística parece envolver três problemas distintos: a origem das variações linguísticas; a difusão e propagação das mudanças linguísticas; e a regularidade da mudança linguística”. Há três

princípios teóricos fundamentais que permeiam a questão central da mudança linguística: a. se uma comunidade heterogênea e plural faz uso de um sistema linguístico, este também deve ser heterogêneo e plural; b. os processos de mudança de uma comunidade de fala se atualizam na variação em cada momento nos padrões de comportamento linguístico observados nessa comunidade; c. se uma mudança implica necessariamente variação, a variação não implica necessariamente mudança (LABOV, 1994; WEINREICH; LABOV; HERZOG, 2006 [1968]).

Ao desenvolver uma pesquisa sociolinguística direcionada pelo aporte teórico-metodológico proposto por Labov (1994 [1972]), é preciso definir os conceitos que permeiam essa corrente de estudos: variação linguística, variantes, variáveis dependentes e variáveis independentes ou explanatórias (linguísticas e sociais).

Segundo Weinreich, Labov e Herzog (2006 [1968]), o termo “variação linguística” é utilizado para se referir às diversas formas linguísticas ou formas alternativas de se dizer a mesma coisa em um mesmo contexto e com o mesmo valor de verdade. A essas formas em variação dá-se o nome de “variantes linguísticas” e ao conjunto dessas variantes dá-se o nome de “variável linguística”. Os conjuntos de variáveis subdividem-se em variáveis linguísticas dependentes e independentes, cuja quantidade vai variar de acordo com a necessidade de cada pesquisa sociolinguística. A “variável dependente” é o foco do estudo, isto é, o fenômeno linguístico que se objetiva estudar. As variantes são os fatores dentro de cada variável considerada, por exemplo, a animacidade do antecedente (+ animado; - animado). Para compreender o uso de uma ou outra variante, debruça-se sobre as “variáveis explanatórias ou independentes”, que, por sua vez, subdividem-se em linguísticas (estruturais) e extralinguísticas (sociais), também definidas como os fatores condicionantes ou que condicionam determinada variação ou mudança linguística.

Sabe-se que para desenvolver um estudo sociolinguístico é preciso definir uma comunidade de fala. Segundo Labov (1994 [1972], p. 87), uma comunidade de fala é entendida como um grupo de pessoas que compartilham traços linguísticos e que distinguem seu grupo de outros. Não é um grupo que fala exatamente igual; os indivíduos se comunicam relativamente mais entre si do que com os outros e, principalmente, compartilham normas e atitudes diante do uso da linguagem.

Ao considerar uma comunidade de fala e realizar os registros orais e vernaculares da mesma, o pesquisador sociolinguista define um fenômeno para ser analisado. Após escolher o fenômeno, é preciso definir a variável de análise e quais variantes estão (ou estiveram) em co-ocorrência durante sua produção. Em seguida, são delimitados fatores condicionantes para aquela variação ou que podem ter direcionado determinada mudança linguística. Realizadas as transcrições, o pesquisador organiza um arquivo de codificação com todas as ocorrências do fenômeno, codificação essa que pode variar a depender do tipo de programa a ser usado (GoldVarb X ou R), para, então, proceder à análise quantitativa dos dados. Explana-se, a seguir, a metodologia quantitativa utilizada nos estudos sociolinguísticos.

2.2 A METODOLOGIA QUANTITATIVA NOS ESTUDOS SOCIOLINGÜÍSTICOS

Há cinquenta anos que Labov (1966) e seus colegas propuseram a análise de regra variáveis para o estudo do fenômeno linguístico. Desde então, diversos estudos têm demonstrado a importância do modelo quantitativo para a variação de dados, seja em dados escritos históricos ou em *corpora* de larga escala. Os estudos sociolinguísticos fazem uso de uma metodologia quantitativa e têm como ferramenta de análise os programas de regra variável e testes de significância. O programa de regra variável é uma ferramenta que conduz análises estatísticas sofisticadas e ajuda a compreender e organizar os dados linguísticos, já que sua enorme produtividade não permite uma análise manual.

O que atualmente se chama de “regra variável” está correlacionado à noção de “heterogeneidade ordenada”, isto é, a variação linguística não é aleatória nem independente, mas sistemática e governada por regras (WEINREICH; LABOV; HERZOG, 2006 [1968], p. 100). As regras variáveis foram inicialmente introduzidas por Labov (1966), quando observou que os falantes fazem escolhas ao usar a linguagem e que essas escolhas são sistemáticas. Devido a essa sistematicidade, a frequência relativa dessa seleção pode ser predita. A regra variável foi projetada como um modelo empírico e explicativo para o fenômeno da fala espontânea, introduzindo um componente probabilístico nos estudos linguísticos.

Segundo Guy e Zilles (2007, p. 26), uma análise quantitativa possui três fases principais: (a) coleta de dados, que objetiva constituir uma amostra representativa possível da comunidade de fala que se deseja observar, o que é mais facilmente obtido através da coleta aleatória; (b) redução dos dados, de modo a garantir um panorama geral do fenômeno, sem enviesar os dados originais, mas minimizando detalhes insignificantes, o que é eficazmente conseguido através de ferramentas estatísticas, como as medidas de tendência (média, mediana e moda), as medidas comparáveis (proporção ou porcentagem) e os índices e, assim também, através da sua posterior apresentação, com o uso de gráficos e tabelas; e, finalmente, (c) a interpretação dos dados, momento em que o linguista parte para a análise do fenômeno em estudo.

O autor indica o uso de dois métodos estatísticos que são comumente usados em estudos sociolinguísticos: o teste de significância, que testa a probabilidade de a “hipótese nula” ser verdadeira, e a análise de regra variável, que se presta a identificar os fatores contextuais influentes em uma variável linguística e a verificar sua relevância.

O programa computacional que, inicialmente, oferece a análise estatística dos registros de fala espontânea é o pacote GoldVarb X. Hoje em dia existem outros programas computacionais estatísticos para análise linguística, como por exemplo o Rbrul e o R (ou RStudio). O GoldVarb X possibilita conduzir dois tipos de análises de dados de variação: 1) Rodada primária binominal (*binominal step-up*) e 2) Rodada secundária binominal (*binominal step-up/step-down*). A rodada primária analisa todos os grupos de fatores e variáveis ao mesmo tempo, permitindo interpretar quais as combinações possíveis, com a finalidade de entender a distribuição geral dos dados. A rodada secundária possibilita diversas rodadas computacionais analisando cada grupo individualmente e fornece as “três linhas da evidência”: significância estatística, peso

relativo, classificação dos fatores condicionantes. Essa segunda rodada tem o objetivo de indicar o que é mais relevante dentre o conjunto de X grupos de fatores, isto é, delimita quais os fatores significativos na aplicação da regra variável (ou de um fenômeno linguístico específico). Nessa rodada, todos os grupos de fatores são testados para determinar qual deles aumenta significativamente o nível de probabilidade.

Para compreender o favorecimento ou desfavorecimento de uma variável, é preciso rodar o peso relativo das mesmas. São eles que indicam se a aplicação da regra variável é favorável ou não em determinado grupo de fatores e quais grupos de fatores são ou não favoráveis a essa aplicação. Os pesos relativos dos fatores obtidos com uma análise multivariada são valores que variam de 0 a 1. Quanto mais próximo o peso relativo for do valor 1, mais favorável será a aplicação. Por outro lado, quanto mais próximo de 0 estiver o peso relativo, mais desfavorável será a aplicação. Para alguns teóricos, valores acima de 0,50 são favoráveis à aplicação, enquanto que valores abaixo de 0,50 são desfavoráveis. O peso relativo procura responder o quanto provável é a aplicação da regra variável em determinado contexto.

Tagliamonte (2006) afirma que há dois objetivos ao encontrar a melhor análise para os dados. Por um lado, deve-se encontrar o modelo que melhor se ajuste aos dados. Isso significa, em parte, combinar fatores que não diferem significativamente entre si. Por outro lado, é preciso explicar e demonstrar como a variação está encaixada no subsistema da gramática bem como na comunidade. Ao utilizar o modelo probabilístico na teoria sociolinguística, objetiva-se mostrar que certas variáveis internas ou externas seguem um padrão similar e quais suas implicações para o fenômeno estudado. Para a autora, a vantagem da metodologia quantitativa está fixada na sua capacidade de modelar estatisticamente os fatores simultâneos e multidimensionais que impactam e condicionam as escolhas dos falantes, para identificar nos dados tendências e regularidades gramaticais até mesmo sutis e avaliar sua produtividade e significância relativas. Essas medidas estatísticas fornecem a base para a pesquisa sociolinguística. A autora alerta, porém, que essas técnicas sofisticadas só são boas mediante bons procedimentos analíticos e metodológicos.

Em consonância com Tagliamonte (2006), o método quantitativo é o aspecto mais importante da análise variacionista, senão o mais importante; é, com certeza, o mais revolucionário, já que mudou e facilitou o modo como se faz linguística nos dias atuais:

Talvez o aspecto mais importante da análise da variação que a diferencia da maioria das outras áreas da linguística, e mesmo da sociolinguística, seja sua abordagem quantitativa. A combinação de técnicas empregadas na análise da variação faz parte da vertente “descritivo-interpretativa” da pesquisa linguística moderna (TAGLIAMONTE, 2006, p. 12)³.

³ Do original: “Perhaps the most important aspect of variation analysis that sets it apart from most other areas of linguistics, and even sociolinguistics, is its quantitative approach. The combination of techniques employed in variation analysis forms part of the ‘descriptive-interpretative’ strand of modern linguistic research” (tradução nossa).

O objetivo essencial da análise da variação é entender o comportamento da variável dependente de acordo com uma série de fatores, externos (sociais) ou internos (linguísticos), ou seja, os contextos multidimensionais em que ocorre.

Vale ressaltar que, embora os programas sejam instrumentos estatísticos de bastante relevância para qualquer sociolinguista, eles não devem ser configurados como único meio de análise. Os dados percentuais, os pesos relativos e os testes de significância ou qui-quadrado nada significam sem uma fiel e adequada interpretação linguística dos resultados encontrados. Concorda-se com Guy (1993, p. 243), ao afirmar que o objetivo primordial de qualquer estudo quantitativo não é produzir números ou prover um resumo estatístico, mas identificar e explicar os fenômenos linguísticos.

3 AS COMUNIDADES DE FALA E A SUA SÓCIO-HISTÓRIA

Nesta seção, são descritas, de modo conciso, a região do Oeste baiano e as duas comunidades de fala da pesquisa que estão nela localizadas: SAMAVI e Montevidinha. A cidade de SAMAVI encontra-se a 860km da capital Salvador e contém uma extensa área rural, abastecida pelo Rio Corrente, que corta a cidade e deságua no Rio São Francisco, enquanto Montevidinha faz parte de um agrupamento de comunidades reconhecidas como remanescentes quilombolas, registrando aproximadamente 200 habitantes, e configura-se como área rural do Município de SAMAVI.

São pontuadas também algumas informações a respeito da sócio-história da cidade, de sua formação e de sua constituição emancipatória e populacional, com base no acervo bibliográfico e nas conclusões parciais do projeto “O português rural do oeste da Bahia”, vinculado ao projeto “Falares do Além São Francisco”. Descreve-se brevemente a comunidade de Montevidinha e seu aspecto remanescente quilombola, o que pode fornecer subsídios para discussões a respeito do contato linguístico na Bahia e, conseqüentemente, no Brasil, dando margens para um maior entendimento acerca do PB.

3.1 O OESTE BAIANO DO ALÉM SÃO FRANCISCO

Como o próprio título sugere, o Oeste baiano do além São Francisco é toda a região do extremo oeste da Bahia que está localizada à margem esquerda do Rio São Francisco. Segundo o IBGE, O Oeste baiano configura todo o território da Mesorregião do Extremo Oeste e parte dos municípios da Mesorregião do Vale do São Francisco e designa todo o território à esquerda do Rio São Francisco, onde estão localizados 35 municípios baianos, em uma área de pouco mais de 183 mil km (BRANDÃO, 2009, p. 48), como se pode visualizar na Figura 1.

Figura 1 – Localização da região do Oeste da Bahia – Brasil



Fonte: Google Imagens.

A formação sócio-histórica da região, assim como de toda a Bahia e do Brasil afora, tem sua origem no processo de ocupação lusitana, relacionado à apropriação do território brasileiro. Para Freire (1998), o contexto de tomada do território e colonização portuguesa no Oeste baiano foram marcados por quatro acontecimentos: (i) a descoberta do Rio São Francisco; (ii) a doação de terras no regime de sesmarias; (iii) a instalação das primeiras fazendas de gado; e (iv) a exploração/submissão da mão de obra indígena, resultando em um quase total extermínio das populações locais. Com base em Oliveira (2015), a formação territorial do Oeste da Bahia também foi resultado das incursões portuguesas, na busca pela exploração dos recursos naturais até então intocados ou salvaguardados pelos nativos que vivem nas margens do Rio São Francisco. A autora, ao referenciar Freire (1998), afirma que

A princípio, o movimento de interiorização compreendido entre os séculos XVI e XVII, com as incursões portuguesas, se davam na direção do interior (de leste ao oeste), representando para os colonizadores um movimento de captura e batalhas contra os indígenas, que naquele contexto estavam empenhados em dizimar as populações nativas, devido sobretudo, às constantes “invasões” dos índios às primeiras nucleações no litoral. Num segundo momento, os interesses econômicos – especialmente na lavoura açucareira, criação de gado e exploração de minérios –, somados à difusão da doutrina católica, tornaram-se as principais causas das explorações e entradas pelo “sertão”, grande parte delas culminando nas margens do Rio São Francisco (OLIVEIRA, 2015, p. 42).

Pode-se dizer que o Rio São Francisco foi o grande responsável pela constituição histórica e populacional da região, pois, por possuir um clima seco e árido, forçava os nativos a se abrigarem no entorno de suas margens, para possibilitar o cultivo e a sobrevivência. Para Rocha (2004, p. 114), “o Rio São Francisco, em meados do século

XVI, já começara a representar um fator importante de ocupação do espaço, posteriormente sendo considerado a principal via de ligação e abastecimento entre as minas de ouro do século XVIII”.

O Oeste baiano começa a se urbanizar com o surgimento de nucleações ou agrupamentos de pessoas ao redor dos afluentes do Rio São Francisco, sendo um deles o Rio Corrente, onde surgiram as cidades de Correntina, Santa Maria da Vitória e São Félix do Coribe. Com a ascensão de antigos arraiais à condição de vilas e cidades, foi sendo consolidada uma rede urbana na região que tomava toda a margem esquerda do Rio São Francisco. Brandão (2010) esclarece que

[...] ainda que os primeiros municípios criados no Além São Francisco sob jurisdição baiana tenham sido Campo Largo (1820), atual distrito de Taguá, em Cotegipe, Carinhanha (1832) e Santa Rita de Cássia (1840), então chamada de Santa Rita do Rio Preto, todos desmembrados de Barra, foi em finais do século XIX que houve maior profusão de novas municipalidades, entre as quais a Vila de Nossa Senhora da Glória do Rio das Éguas, hoje Correntina (1866), Porto de Santa Maria da Vitória do Rio Corrente (1880), cuja denominação atual é apenas Santa Maria da Vitória, Angical (então conhecida como Santana de Angical), Casa Nova (1879) e Santana (1890), além de Barreiras (1891), redefinindo significativamente a organização do espaço intra-regional e dando vazão aos desejos políticos das lideranças locais (BRANDÃO, 2010, p. 44).

De acordo com registros bibliográficos, antes da independência, essa região já vivenciava diversos contatos e conflitos; essas disputas não estavam apenas relacionadas à posse da grande vastidão de terra mas também ao próprio contato entre europeus e indígenas. Posteriormente, escravos, africanos ou nascidos no Brasil, passaram a fazer parte da composição social da região.

Embora seja escassa a literatura, há registros – escritos e orais (depoimentos) – de que o Oeste baiano abrigou indígenas e africanos fugidos. Sobre a presença indígena na região, esta pode ser encontrada ao longo de todo o Vale do São Francisco e no chamado Além São Francisco, por meio de registros dos antepassados indígenas como os grafismos e pinturas rupestres da Lapa dos Tapuias ou Gruta das Pedras Brillhantes, no município de São Desidério (GALVÃO et al., 2012, p. 25). Pierson (1972, p. 228-229) destaca que cerca de nove grupos indígenas viveram ao longo do Médio São Francisco: Tamoio e Cataguá, Shacriaba, Acroá, Aricobé, Tobajara, Amoipira, Tupiná, Ocren e Sacragrinha e Tupinambá, além dos Caiapós, mencionados anteriormente. Sobre os africanos, Souza (1996) encontrou alta porcentagem de sua presença na história da região.

Assim como em todo o Vale do São Francisco, as comunidades de SAMAVI e Montevidinha (esta reconhecidamente quilombola) se apresentam com características indígenas, africanas e europeias (SOUZA, 1996, p. 34) e configuram-se como representativas do contexto sócio-histórico do Brasil.

Esses indícios da presença indígena e africana na região validam a importância de estudar os fatos linguísticos e o contato na formação do português do Oeste baiano, seja pela necessidade de pesquisa linguística na região, seja pela importância de documentar e registrar pesquisas sobre a região. É importante mencionar que ainda há muito o que

estudar sobre a história do Oeste da Bahia e sua formação sociocultural, como constata Neves (2012):

Por se dedicar durante muito tempo, especialmente a Salvador e ao Recôncavo, a pesquisa histórica na Bahia pouco produziu sobre o semiárido, durante a colonização portuguesa [...]. No império quase se ignorou a história dos sertões e na Primeira República focalizou-se pouco mais que especificidades temáticas como Canudos, cangaço e o arbítrio político de coronéis da Guarda Nacional, no exército, poderes locais, sempre com vagas informações sobre a região do Oeste da Bahia ou nenhuma alusão a ela (NEVES, 2012, p. 33).

Como já apontam Figueiredo (2015) e Barros (2015), o entendimento do funcionamento e da organização sócio-histórica do Oeste da Bahia não se constitui uma tarefa fácil, dado o raro acervo de estudos e pesquisas sobre a sua formação, sobretudo no que se refere à presença dos negros e índios. Entretanto Souza (1996, p. 34) afirma que, como consequência da miscigenação das raças, a cultura, ali no Município de Cocos, como de resto em todo vale do São Francisco, se apresenta com características indígenas, africanas e europeias.

Em relação à presença dos negros, Souza (1996), observando o livro de registro de batismo da freguesia de São José de Carinhanha, datado de 1804 a 1909, percebeu uma alta porcentagem deles em diversas condições sociais (cativos, forros, crioulos e pardos livres) nas regiões de Cocos, Coribe, Feira da Mata, Correntina, Santa Maria da Vitória, Santana, Serra Dourada e Canápolis. Em relação à presença dos índios, o autor diz que, à margem esquerda do Rio São Francisco, viviam os índios caiapós, que, com a chegada das Bandeiras, dirigiram-se para o alto da Caririnha, onde se refugiaram, o que justifica seus traços fenóticos serem bastante peculiares nessas margens. Esses índios caiapós, nos termos do autor, “eram valentes e temíveis” e, com isso, resistiram e lutaram contra os bandeirantes baianos e paulistas. Apenas no final do século XIX é que bateram em retirada à Amazônia. O autor aponta ainda a presença de descendentes de índios foragidos de São João das Missões.

Tais pistas ratificam a importância de estudar os fatos sócio-históricos na formação do português do Oeste baiano, já que, em sua constituição, verifica-se a forte presença de índios, além da presença dos negros, como afirma Souza (1996), o que aponta para uma situação de contato linguístico. Na seção 3.2 é abordada, de modo breve, a sócio-história de SAMAVI e alguns aspectos característicos de Montevidinha.

3.2 A CIDADE DE SANTA MARIA DA VITÓRIA E A COMUNIDADE DE MONTEVIDINHA

Santa Maria da Vitória ou SAMAVI, como é carinhosamente referida por seus moradores, é uma cidade que está situada às margens do Rio Corrente, que se divide entre a cidade de SAMAVI e São Félix do Coribe. O município de Santa Maria da Vitória teve origem, nos meados do século XIX, num arraial formado na margem do Rio Corrente, território até então pertencente ao Município de Rio das Éguas (conhecido atualmente como Correntina). Sua formação urbana iniciou-se com pequenos grupos de pessoas que

para ali foram com a finalidade de exploração de ouro nas proximidades. Posteriormente, ao se instalar, essa população passou a dedicar-se à agricultura.

Em 1808, a localidade categorizava-se como uma vila, conhecida como Porto de Santa Maria da Vitória do Rio Corrente, já que ali havia uma espécie de porto em que atracavam as embarcações de cargas para comercialização, venda e troca de produtos entre os viajantes e os nativos da cidade. Encontravam-se na vila poucas casas, circundadas de frondosas gameleiras, em cuja sombra se abrigavam os que vinha fazer transações comerciais. Em 1850, um artífice, vindo da cidade de Barra do Rio Grande, construiu a primeira embarcação para o transporte de mercadorias e animais. Foram construídas logo após outras embarcações e o arraial começou a crescer com a chegada de grande número de pessoas para as atividades agrícolas. Em 26 de junho de 1909, a Vila de Santa Maria foi elevada à condição de cidade pela Lei nº 737, porém com a denominação de Santa Maria. Foi somente em 1944 que o município passou a ser denominado Santa Maria da Vitória.

Segundo o IBGE (2017), a população atual do município de Santa Maria da Vitória é de 41.769 habitantes. É uma das principais cidades do Oeste da Bahia e a principal cidade da Bacia do Rio Corrente, que é composta 11 municípios: Brejolândia, Canápolis, Cocos, Coribe, Correntina, Jaborandi, Santana, Santa Maria da Vitória, São Felix do Coribe, Serra Dourada e Tabocas do Brejo Velho. A cidade está localizada a 866 km de Salvador e a 220 km de Barreiras. Na Figura 2, a seguir, é possível visualizar a localização de SAMAVI no Estado da Bahia.

Figura 2 – Localização de Santa Maria da Vitória – Brasil



Fonte: Google Imagens.

O entorno da cidade, assim como o de milhares de cidades brasileiras, é cercado por povoados ou pequenos agrupamentos de comunidades. Montevidinha é uma das maiores comunidades rurais da cidade e faz parte de um agrupamento de comunidades reconhecidas como remanescentes quilombolas, registrando aproximadamente 200 habitantes e configura-se como área rural do município de SAMAVI. De acordo com Bello, Oliveira e Oliveira (2010), Montevidinha está localizada na zona rural de SAMAVI, insere-se no contexto de comunidade remanescente de antigos escravos.

4 A METODOLOGIA: CONSTITUIÇÃO DE *CORPUS*

Segundo Oushiro (2014, p. 134), ao lidar com uma grande quantidade de dados, o pesquisador sociolinguista deve realizar os seguintes passos: 1) a coleta de dados com gravações de entrevistas; 2) a transcrição dessas gravações; 3) a definição de uma variável sociolinguística e de seus contextos (sócio)linguísticos possíveis; 4) a identificação de ocorrências no *corpus* de entrevistas; 5) o levantamento de hipóteses sobre fatores, de natureza social e linguística, que estejam correlacionados ao uso da variável; 6) a codificação das ocorrências de acordo com as hipóteses levantadas; 7) a análise quantitativa dos dados no programa estatístico; e 8) a interpretação dos resultados obtidos.

Nesta seção, apresenta-se a metodologia para o levantamento dos dados linguísticos utilizados como objeto da pesquisa. Os critérios para a seleção dos informantes seguiram as especificações do projeto “Falares baianos do Além São Francisco”⁴ (FIGUEIREDO, 2015) e foram realizadas entrevistas com falantes nativos ou que tenham se estabelecido em Santa Maria da Vitória e Montevidinha até os cinco anos de idade e residam lá desde então.

Para a realização das entrevistas, optou-se pelo tipo de entrevista DID (Diálogo entre Informante e Documentador), com o objetivo de registrar dados de fala espontânea, abordando temas de interesse da comunidade, como o cotidiano da cidade, a infância dos informantes, suas profissões, gostos musicais e cinematográficos, sua rotina, histórias da cidade, dentre outros assuntos.

Com vistas a fornecer um ambiente confortável ao informante e a permitir o prolongamento das narrativas, prezou-se pelo cuidado de amenizar o paradoxo do observador (LABOV, 1994 [1972]). Durante o desenvolvimento da coleta desses dados, as gravações foram conduzidas com auxílio de gravadores e/ou computadores com placa de som e tiveram entre 45 minutos e uma hora e 10 minutos de duração.

Antes da seleção dos informantes, fez-se um levantamento dos mesmos utilizando uma célula em que constavam informações relevantes para a pesquisa e que considerava as três estratificações sociais de base para todo e qualquer trabalho sociolinguístico: sexo, faixa etária e escolaridade. Ao finalizar o levantamento dos informantes, contactou-se cada um deles individualmente para a realização das entrevistas.

Para a coleta e constituição da amostra de SAMAVI, foram selecionados 12 informantes (apresentados no Quadro 1), de acordo com as seguintes variáveis sociais: (i) Sexo: oito homens e oito mulheres; (ii) Faixa etária: I – de 25 a 35 anos, II – de 45 a 55 anos, III – acima de 65 anos; (iii) Escolaridade: baixa – analfabetos e com pouco acesso à escola (pelo menos quatro anos) e alta – ensino médio completo⁵ ou com ensino universitário.

⁴ O projeto “Falares baianos do Além São Francisco” foi idealizado pela Profa. Maria Cristina Figueiredo e pela Profa. Ísis Juliana Figueiredo de Barros, a fim de criar um acervo de amostras de fala de comunidades do Oeste baiano e investigar os fenômenos linguísticos variáveis, caracterizadores do português falado na região, contribuindo, assim, para a compreensão do PB. A pesquisa propõe ampliar o mapeamento sociolinguístico do Estado da Bahia, contribuindo para a construção histórica do PB.

⁵ Considerou-se ensino médio completo como alta escolaridade pela dificuldade em encontrar informantes com ensino universitário completo na cidade, já que as instituições de ensino superior são recentes na região.

Quadro 1: Informantes da amostra de fala de Santa Maria da Vitória

FAIXA ETÁRIA	SEXO			
	Mulher	Mulher	Homem	Homem
Faixa I (25 a 35 anos)	INF01 – 32 anos, baixa escolaridade, doméstica	INF02 – 27 anos, alta escolaridade, gerente	INF03 – 34 anos, baixa escolaridade, segurança (autônomo)	INF04 – 30 anos, alta escolaridade, empresário (dono de loja)
Faixa II (45 a 55 anos)	INF05 – 54 anos, baixa escolaridade, lavradora	INF06 – 54 anos, alta escolaridade, cozinheira	INF07 – 49 anos, baixa escolaridade, pescador	INF08 – 46 anos, alta escolaridade, comerciante
Faixa III (acima de 65 anos)	INF09 – 88 anos, baixa escolaridade, costureira	INF10 – 68 anos, alta escolaridade, professora aposentada	INF11 – 74 anos, baixa escolaridade, músico e sapateiro	INF12 – 77 anos, alta escolaridade, advogado aposentado

Fonte: Elaborado pelos autores.

Findada a realização das entrevistas, foi realizado um levantamento bibliográfico sobre a sócio-história da cidade em livros e documentos de acervos públicos ou pessoais, além da gravação de depoimentos. A pesquisa prossegue com as transcrições dos inquéritos, a partir dos áudios em extensão .mp3, com o auxílio do programa *Transcriber*, utilizando a chave de transcrição proposta no projeto “Vertentes do português popular do Estado da Bahia”, fundado pelo Prof. Dr. Dante Lucchesi e coordenado pelo Prof. Dr. Gredson dos Santos.

Finalizadas as transcrições, serão definidas as variáveis sociolinguísticas e a identificação de suas ocorrências no *corpus*, seguidas do levantamento de hipóteses sobre fatores condicionantes de natureza social e linguística que estejam correlacionados ao uso da variável e que possam explicá-lo.

Após esses procedimentos, será feita a codificação das ocorrências de acordo com as hipóteses levantadas e com uma chave de codificação que será definida posteriormente. Em seguida, será realizada a análise quantitativa dos dados no GoldVarb X ou no R para o tratamento estatístico e probabilístico das ocorrências em cruzamento com os grupos de fatores das variáveis. Por fim, chegará a etapa de interpretação e divulgação dos resultados obtidos. Vale ressaltar que “o trabalho quantitativo não é um substituto, mas um acessório para a análise linguística [...] Ele não nos diz o que os números significam, muito menos faz linguística por nós [...] a resposta vem de nossa teoria linguística, não de um programa estatístico” (GUY; ZILLES, 2007, p. 65). É, portanto, função do pesquisador interpretar os resultados encontrados, confrontando a literatura linguística com os encaixamentos sociais e linguísticos de determinado fenômeno naquela comunidade de fala.

5 CONSIDERAÇÕES FINAIS

Espera-se que o estudo contribua para a compreensão e descrição linguísticas dos falares baianos do Oeste da Bahia, da sua formação sócio-histórica, através da constituição de *corpus* em Santa Maria da Vitória e Montevidinha e contribua também para a compreensão da formação sócio-histórica do PB e para ampliação dos estudos sociolinguísticos no Brasil, fornecendo subsídios a favor da hipótese da transmissão linguística irregular para explicar a formação do PB como resultado do contato linguístico.

Além disso, ao investigar o uso variável das estratégias de relativização (especificamente para esta pesquisa) e de outros fenômenos possíveis, pretende-se caracterizar e descrever o português falado no Oeste Baiano, colaborando, assim, para uma maior descrição do PB. Ao realizar a análise e interpretação dos dados, espera-se ser possível delimitar quais fatores linguísticos e/ou sociais condicionam e/ou favorecem a realização das estratégias de relativização no português do Oeste baiano. Com o mapeamento do uso das relativas na Bahia, o resultado que se pretende alcançar é um estudo comparativo, descrevendo os dados e confrontando-os com trabalhos realizados sobre o tema nos diversos falares e localidades do estado baiano, moldando um recorte do sistema sintático do português baiano no que diz respeito à relativização e organizando publicações de artigo e/ou livro, contribuindo para a produção científica no estado e no país.

Não menos importante, almeja-se reunir informações bibliográficas e demográficas sobre a sócio-história da região, a fim de compreender os fatos extralinguísticos inerentes à fala das comunidades, a partir da leitura dos poucos trabalhos já realizados, colaborando para uma ampliação do acervo documental sobre a região, divulgando características singulares dos municípios e de seus residentes. Por fim, acredita-se que, com a reunião de depoimentos dos informantes e dos registros junto aos órgãos públicos, será possível materializar a memória histórica e cultural de SAMAVI e Montevidinha, promovendo a ciência, a construção identitária e a valorização de um povo, a partir da divulgação dos falares e da cultura do interior do Oeste da Bahia.

REFERÊNCIAS

ARAÚJO, S. S. F. O português popular do semiárido baiano: fundamentos teóricos, sócio-históricos e empíricos. In: LOPES, N. S.; OLIVEIRA, J. M.; PARCERO, L. M. J. *Estudos sobre o português do Nordeste: língua, lugar e sociedade*. São Paulo: Blucher, 2017. p. 45-72.

BAXTER, A. N. Transmissão geracional irregular na história do português brasileiro – divergências nas vertentes afro-brasileiras. *Revista Internacional de Língua Portuguesa*. Lisboa, n. 14, p. 72-90, 1995.

- BELLO, J. M. F. O.; OLIVEIRA, L. S. F. C.; OLIVEIRA, R. T. *A comunidade Montevidinba no contexto de reconhecimento como quilombola no Oeste Baiano (1870-2010)*. 2010. Monografia. Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2010.
- BRANDÃO, P. R. B. Um território indiferenciado dos Sertões: a geografia pretérita do Oeste baiano (1501-1827). *Boletim Goiano de Geografia*. Goiânia, v. 29, n. 1, p. 47-56, 2009.
- BRANDÃO, P. R. B. A formação territorial do Oeste baiano: a constituição do “Além São Francisco” (1827 – 1985). *Geotextos*. Salvador, v. 6, n. 1, p. 35-50, jul./2010.
- BARROS, Isis J. F. O português rural do oeste da Bahia. 2015. 15 f. (Projeto de pesquisa) - Centro Multidisciplinar de Santa Maria da Vitória, UFOB, Santa Maria da Vitória, 2015.
- CARDOSO, S. A. M. *Geolinguística: tradição e modernidade*. São Paulo: Parábola, 2010.
- FREIRE, F. *História territorial do Brasil*. Salvador: Secretaria da Cultura e Turismo; Instituto Geográfico e Histórico da Bahia, 1998.
- FREITAG, R. M. Ko. *Metodologia de coleta e manipulação de dados em sociolinguística*. São Paulo: Blucher, 2014.
- FIGUEIREDO, Maria Cristina. Os Falares do Além São Francisco. 2015. 25f. (Projeto de Pesquisa) - Instituto de Letras, UFBA, Salvador, 2015.
- FIGUEIREDO, C. *Falares baianos do além São Francisco*. Salvador: UFBA, 2015.
- GALVÃO, A. L. O. et al. Breve descrição do patrimônio espeleológico do Município de São Desidério – BA. *Revista Brasileira de Espeleologia*. Brasília, v. 2, n. 1, 2012, p. 13-28.
- GUY, G. R. The quantitative analysis of linguistic variation. In: PRESTON, D. (Ed.). *American dialect research*. Amsterdam: John Benjamins, 1993. p. 223-249.
- GUY, G. R.; ZILLES, A. *Sociolinguística quantitativa: instrumental de análise*. São Paulo: Parábola, 2007.
- IBGE. *Instituto Brasileiro de Geografia e Estatística*. 2017. Disponível em: <<https://cidades.ibge.gov.br/brasil/ba/santa-maria-da-vitoria>>. Acesso em: 24 ago. 2018.
- LABOV, W. *The social stratification of English in New York City*. Washington, DC: Center for Applied Linguistics, 1966.

- LABOV, W. *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press, 1994 [1972].
- LUCCHESI, D; BAXTER, A. A transmissão linguística irregular. In: LUCCHESI, D.; BAXTER, A.; RIBEIRO, I. (Orgs.). *O português afro-brasileiro*. Salvador: EDUFBA, 2009. p. 101-124.
- NEVES, E. F. Propriedade, posse e exploração da terra: domínio fundiário na Região Oeste da Bahia, século XIX. In: CARIBÉ, C.; VALE, R. (Orgs.). *Oeste da Bahia: trilhando velhos e novos caminhos do além São Francisco*. Feira de Santana: UEFS, 2012. p. 31-96.
- OLIVEIRA, M. A. B. Abordagem histórico-geográfica do Oeste da Bahia: um breve levantamento dos movimentos de divisão no contexto de formação territorial. *Revista Espaço Acadêmico*. Maringá, n. 165, p. 41-53, fev./2015.
- OUSHIRO, L. Tratamento de dados com o R para análises sociolinguísticas. In: FREITAG, R. M. Ko. *Metodologia de coleta e manipulação de dados em sociolinguística*. São Paulo: Blucher, 2014. p. 133-176.
- PIERSON, D. *O homem no Vale do São Francisco*. Trad. Maria Aparecida Kerberg e Ruy Jungmann. Tomo I. Rio de Janeiro: SUVALE, 1972.
- ROCHA, G. *O Rio São Francisco: fator precípua da existência do Brasil*. 4. ed. São Paulo: Companhia Editora Nacional, 2004.
- SOUZA, J. E. *O sertão de cocos na Bahia: uma miragem no Oeste*. Brasília: Arte e Movimento, 1996.
- TAGLIAMONTE, S. *Analysing sociolinguistic variation*. Cambridge: Cambridge University Press, 2006.
- WEINREICH, U. LABOV, W. HERZOG, M. *Fundamentos empíricos para uma teoria da mudança linguística*. Trad. Marcos Bagno; revisão técnica de Carlos Alberto Faraco; posfácio de Maria da Conceição Paiva e Maria Eugênia L. Duarte. São Paulo: Parábola, 2006 [1968].