

O NÚCLEO DE ESTUDOS DA LÍNGUA PORTUGUESA E OS ESTUDOS SOBRE A FORMAÇÃO DA REALIDADE SOCIOLINGÜÍSTICA BRASILEIRA¹

*Silvana Silva de Farias Araujo
Zenaide de Oliveira Novais Carneiro
Mariana Fagundes de Oliveira Lacerda*

Neste texto, discutimos questões sobre a identidade linguística brasileira, tecendo algumas considerações sobre a formação do português brasileiro (PB) e sua caracterização atual. Ao longo do texto, fica evidente que, ao se discutir essas questões, a agenda de pesquisa deve estar voltada para a vertente popular dessa variedade linguística, pois esta traz, de maneira mais sistematizada, evidências empíricas que apontam para a diferenciação entre o PB e o português europeu (PE), sobressaindo o vernáculo brasileiro. Será evidenciado também que, não obstante a importância de estudos sobre o português popular brasileiro, pesquisas que tenham como foco a vertente culta do PB não devam ser desconsideradas.

A discussão está embasada na hipótese de que o contato linguístico entre portugueses de diferentes estratos sociais e de diversas

¹ As autoras dedicam este texto às suas colegas e amigas Norma Lucia Fernandes de Almeida e Eliana Pitombo Teixeira, parceiras inseparáveis de pesquisa que, por especificidades desta obra, não puderam ser coautoras deste capítulo.

regiões de Portugal, indígenas de diversas etnias e africanos falantes nativos de diversas línguas africanas foi um fator decisivo para a formação de uma variedade linguística distinta da lusitana. Assume-se, portanto, que esse contato entre línguas diversas e ininteligíveis, que levou à aquisição do português como segunda língua por parte dos aloglotas no Brasil, tem reflexos na realidade linguística até os dias atuais.

Além dessas questões históricas e de algumas sobre o PB atual, discutimos uma terceira neste texto, a saber, as contribuições do Núcleo de Estudos da Língua Portuguesa (Nelp), sediado no Departamento de Letras e Artes (DLA) da Universidade Estadual de Feira de Santana (UEFS), para os estudos sobre formação e caracterização da realidade sociolinguística brasileira.

BREVES CONSIDERAÇÕES SOBRE A FORMAÇÃO DO PORTUGUÊS BRASILEIRO

Historicamente, a nação brasileira sempre foi marcada por sua plurietnia, haja vista que, ao concretizarem o seu projeto colonizador, os portugueses conviveram com diversos grupos indígenas (que falavam cerca de mil línguas) e com cerca de quatro milhões de africanos (que falavam cerca de duzentas línguas), conforme quantificação divulgada por Lucchesi (2009, p. 41). Sendo assim, o contato linguístico e cultural esteve presente na sócio-história brasileira, levando Houaiss (1985, p. 91) a firmar que “o PB nasce com diversidade”.

Nesse sentido, frustradas as primeiras tentativas de emprego do braço indígena, deu-se ensejo, no Brasil, à introdução de escravos africanos. A participação africana na colonização brasileira foi mais acentuada do que a indígena, sendo consensual entre os estudiosos

Diálogos Linguísticos e Filológicos

variação, cultura escrita e léxico

a crença de que eles exerceram maior influência para a formação da realidade sociolinguística brasileira. Contudo, não se pode acreditar que a presença de numerosos índios no Brasil não tenha exercido influências na formação do PB.

Sobre a importância de se considerar também a relevância da população indígena para a formação da realidade sociolinguística brasileira, Araujo (2017, p. 50) faz importantes observações:

Nesse sentido, sobressai 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. Por outro lado, fatos da sócio-história do Brasil conduzem a que a população africana seja vista como a grande difusora do “português geral brasileiro”. Entre esses fatos, destaca-se, primeiramente, a intensa participação de negros africanos e de seus descendentes na demografia histórica do Brasil em comparação com a extrema redução dos índios, conforme é evidenciado em alguns estudos, a exemplo do célebre trabalho de Mussa (1991).

Mussa (1991, p. 163), considerando dados de diversas fontes, demonstra que seria plausível que apenas 30% da população brasileira tivesse o domínio da língua portuguesa como língua nativa nos primeiros quatro séculos de formação do país, conforme é evidenciado na Tabela 1.

Tabela 1 – Demografia histórica da sociedade brasileira, segundo proposta de Mussa (1991)

	1538- 1600	1601- 1700	1701- 1800	1801- 1850	1851- 1890
Africanos	20%	30%	20%	12%	2%
Negros brasileiros	-	20%	21%	19%	13%
Mulatos	-	10%	19%	34%	42%
Branco brasileiros	-	5%	10%	17%	24%
Europeus	30%	25%	22%	14%	17%
Índios integrados	50%	10%	8%	4%	2%

Fonte: Mussa (1991, p. 163).

Assim, embora não ignore a participação indígena na difusão do português modificado na sócio-história do PB, Mattos e Silva (2001) traz informações pertinentes para sustentar a hipótese de que a população de origem africana constituiu a força motriz para difundir o português no Brasil em detrimento das línguas indígenas. Dessa maneira, a autora postula que, nos considerados “espaços ilegítimos da escravidão”, a língua utilizada era a portuguesa, e não outras de origem banto, de forma tal que os africanos e seus descendentes difundiram o “português geral brasileiro”. A pesquisadora sustenta, pois, a ideia de que a população africana escravizada no Brasil e os seus descendentes foram os responsáveis por generalizarem os traços linguísticos motivados pelo contato entre línguas na formação do PB.

Nesse cenário de multilinguismo e de trocas culturais diversas, obviamente, não é de surpreender que tenha havido mudanças

linguísticas “de baixo para cima” (LABOV, 2008). Contudo, essas interferências não levaram à formação prototípica de um crioulo, ainda mais porque, no século XVIII, a vinda de muitos portugueses – com o ciclo do ouro e com a transferência da Família Real em 1808 – aumentou a facilidade de acesso à língua-alvo por falantes não nativos do português. Foi marcante para a feição da norma brasileira também a vinda de imigrantes europeus e asiáticos, no final do século XIX e no início do século XX, para trabalharem no Brasil (a princípio, concentrados em estados das regiões Sul e Sudeste e agregados ao estrato baixo da sociedade), os quais também adquiriram a língua portuguesa em condições especiais: como língua segunda e livre de normatizações.

No século XIX, intensifica-se a polarização sociolinguística, presente desde os primeiros anos da colônia, pois a incipiente elite cultural brasileira esforça-se para retirar de cena os matizes que revelassem a ancestralidade da nação, marcada pela ruralidade e pela língua formada por diversos contatos linguísticos. Assim, as características do PB, moldadas desde o início da colonização por processos de transmissão linguística irregular, passaram a sofrer forte tentativa de ocultação por parte daqueles que mantinham fortes elos com a Europa (PAGOTTO, 1998), embora ainda subsistam de forma inequívoca na fala de segmentos populares e até mesmo na fala dos escolarizados, em face das interferências múltiplas.

Desse modo, a língua portuguesa no Brasil, afetada por processos de transmissão linguística irregular, inicialmente na aprendizagem dos índios e, mais tarde, dos negros, foi incorporando fatos linguísticos comuns em situações de contato entre línguas. No decorrer dos séculos, esses fatos foram sendo mais difundidos social e espacialmente. Essa difusão se deve, sobretudo, à questão de europeus (alemães e italianos) e asiáticos generalizarem esses fatos – inclusive no estrato socioeconômico mais alto da sociedade – e pelo

constante deslocamento de escravos nos diversos ciclos econômicos da história brasileira (LUCCHESI, 2009; LUCCHESI; BAXTER, 2009; e MATTOS E SILVA, 2000, 2002, 2008). O projeto da elite brasileira não conseguiu homogeneizar e nem europeizar o PB, mesmo porque outros fatos socioculturais além da demografia histórica, como a escolarização e a urbanização como processos tardios no Brasil, foram imperativos na constituição da realidade sociolinguística brasileira. Esse aspecto é discutido na próxima seção.

REALIDADE SOCIOLINGÜÍSTICA BRASILEIRA ATUAL

No quadro de ampla difusão do português adquirido livre de normatizações, de “oitiva”, como segunda língua, consolidaram-se os matizes que deram o tom do PB, de maneira que se iniciava a polarização sociolinguística do Brasil, principalmente nas concentrações urbanas que já existiam. De um lado, o português “modificado”, falado pelo extenso contingente populacional do Brasil, e de outro, o português idealizado pela escassa elite brasileira, zeladora dos padrões linguísticos e culturais lusitanos.

Assim, a primeira coisa que se deve ter em mente quando se fala em português brasileiro é que esse termo não recobre uma realidade homogênea. Mais ainda: deve-se ter em mente que a realidade linguística é bipolarizada, como bem expôs Lucchesi (1994, 2001, 2015). Certamente, essa realidade atual é fruto das condições de formação do Brasil, com “atores” diversificados.² Assim, como preconceito social e linguístico correlacionam-se, as marcas características das variedades populares brasileiras não demoraram a ser estigmatizadas, uma vez que aqueles que estudaram e/ou que são favoráveis à europeização da cultura brasileira as avaliaram negativamente.

² Mattos e Silva (2009, p. 16) destaca a incomensurabilidade da diversidade de cada um desses atores letrados e não letrados, “que não temos como demonstrar rigorosamente, mas que podemos inferir, aproximadamente”.

Acreditamos que as condições sócio-históricas do PB repercutiram significativamente na estrutura da língua portuguesa, fazendo com que houvesse uma bipolarização de normas linguísticas no Brasil, com um polo que abriga as variedades cultas, essas mais próximas da norma-padrão, e outro que abriga as variedades populares, marcadas por um processo de extrema redução da morfologia flexional. Entendemos, igualmente, que, com as profundas e contínuas mudanças ocorridas no Brasil a partir do século XX, tais normas se entrecruzam, influenciando-se mutuamente (Lucchesi, 2001).

A respeito disso, fica evidente que as variedades populares brasileiras são as que mais preservam os efeitos do contato entre línguas na história sociolinguística do país. Por conseguinte, consideramos que pesquisas com dados do PB não devem desconsiderar a existência de duas histórias sociolinguísticas, realizando análises empíricas de dados com diferentes perfis de maneira conjunta.

A esse respeito, Araujo (2014), ao analisar a concordância verbal com a terceira pessoa do plural na comunidade de fala de Feira de Santana-BA, analisou os dados das duas normas do PB separadamente, além da forma conjunta, de modo que os resultados indicaram que, não obstante as contínuas mudanças ocorridas na dinâmica social do Brasil a partir da década de 1940, com o aumento de estradas e a democratização do acesso ao ensino e aos meios de comunicação de massa – aspectos que aumentam a chance de aquisição de formas linguísticas privilegiadas –, ainda prevalece um hiato entre a fala do segmento culto e a do popular. Em outras palavras, ainda é possível depreender, nos padrões de fala atuais do PB, uma variação diastrática no que tange ao uso da concordância verbal de número, algo que reflete as condições de formação da variedade brasileira da língua portuguesa. Nesse sentido, frisa-se a estreita relação entre baixa ou nula escolaridade e a origem pobre dos informantes, algo que, por sua vez, vincula-se à exclusão a que por séculos foi

submetida uma parcela da população no Brasil, gerando a existência de duas histórias sociolinguísticas brasileiras.

Assim, o PB culto teria como antecedente o português europeu (principalmente das variedades faladas pelos portugueses letrados) e o PB popular teria como antecedente o português geral brasileiro, este em nomenclatura utilizada por Mattos e Silva (2004).

CONTRIBUIÇÃO DO NELP PARA OS ESTUDOS SOBRE A FORMAÇÃO DA REALIDADE SOCIOLINGÜÍSTICA BRASILEIRA

Mattos e Silva (2009, p. 11), ao fornecer uma pauta de estudos sobre a história do português culto e a do popular brasileiro, afirma que, para o primeiro caso, as fontes de arquivo são fundamentais. Já para o segundo, “o caminho será outro, mais espinhoso, mas não impossível”. Assim, nesta seção são abordadas fontes para o estudo diacrônico, sócio-histórico e sociolinguístico dessas duas vertentes.

Para a autora, a reconstrução histórico-diacrônica do passado da norma culta do português brasileiro poderá ser recuperada

[...] pela via tradicional desse tipo de trabalho para as línguas que estão representadas pela escrita, ou seja, pelo escrutínio das fontes documentais remanescentes. (MATTOS E SILVA, 2009, p. 19).

Ela propõe que esse trabalho pode ser feito com base em um *corpus* diacrônico seriado, constituído de documentos de tipos o mais diversificado possível, que possibilitarão a captação das variações linguísticas conviventes que esboçariam prováveis mudanças linguísticas. Ela acrescenta que o percurso para a reconstrução do passado do *português brasileiro popular* não será o mesmo que para a reconstrução do passado do *português brasileiro culto*, que se esteia numa tradição escrita, mas, como foi constituído com base na oralidade, essa via deve

guiar as investigações, o que, para ela, seria um “[...] percurso análogo, *mutatis mutandis* e *modus in rebus*, ao da reconstrução do chamado ‘latim vulgar’” (MATTOS E SILVA, 2009, p. 14).

De forma resumida, podemos afirmar que, para se definir a constituição do português culto brasileiro, faz-se imprescindível a realização de pesquisas com dados de língua escrita do contingente da população escolarizada de tempos pretéritos, quer brasileiros, quer portugueses. Por outro lado, para o escrutínio da constituição histórica das variedades populares do PB, é fundamental a realização de estudos feitos com base em textos escritos por informantes poucos escolarizados, fazendo em ambos os casos – ou seja, tanto nas pesquisas para as faces cultas quanto para as populares – a sócio-história dessas variedades. Outra via de investigação que pode ser realizada é a busca do presente para o passado, por meio de gravações com a fala de pessoas analfabetas ou parcamente escolarizadas, tanto em comunidades urbanas como em comunidades rurais.

Assim, indo na direção das ideias expostas por Mattos e Silva (2009), Lobo (2001, p. 109) discorre sobre a constituição de *corpora* diacrônicos do português brasileiro, definindo que o perfil do que seria um *corpus* linguístico ideal deveria abranger os seguintes subcorpora:

- a) subcorpus 1: as variedades do português europeu transplantadas para o Brasil.
- b) subcorpus 2: as variedades do português falado como segunda língua pelos aloglotas.
- c) subcorpus 3: as variedades do PB que paulatinamente se iam constituindo: (i) as variedades cultas – supostamente mais unitárias e descendentes diretas das variedades do português europeu – e (ii) as variedades populares – supostamente mais diversificadas e descendentes diretas das variedades do português como segunda língua.

Nessa linha de raciocínio, podemos afirmar que as pesquisas que vêm sendo realizadas no Nelp contemplam essas orientações dadas por Mattos e Silva (2009) e Lobo (2001), conforme explanamos em seguida. O Nelp foi criado em 1996, pelas professoras Norma Lucia Fernandes de Almeida e Zenaide de Oliveira Novais, e atualmente congrega diversos projetos de pesquisa do DLA da UEFS. O plano de trabalho, desde a sua fundação, era pautado em formação de banco de dados linguísticos, pesquisas sócio-históricas e análises linguísticas. Os objetivos principais do núcleo são contribuir com o debate sobre a formação do português do Brasil e auxiliar na discussão acerca da realidade sociolinguística brasileira.

Por limitação de espaço, não são tratados todos os projetos de pesquisa alocados no Nelp. Apenas serão abordadas questões sobre três projetos: “A língua portuguesa do semiárido baiano – Fases 1, 2 e 3”; “Vozes do sertão em dados: história, povos e formação do português brasileiro” e “Em busca das raízes do português brasileiro”.

O projeto A Língua Portuguesa no Semiárido Baiano começou a ser executado em 1998, tendo como objetivo geral contribuir para o estudo da sócio-história dessa variedade linguística, na sua modalidade falada pelos segmentos mais desfavorecidos da sociedade, tendo as seguintes frentes de trabalho: a) a formação de banco de dados com amostras de fala de comunidades rurais e urbanas do semiárido baiano; b) a realização de estudos sócio-históricos sobre a formação da língua portuguesa nas localidades pesquisadas; c) a realização de análises linguísticas à luz de diferentes teorias linguísticas; e d) a análise de livros didáticos utilizados nas escolas da região.³

O projeto foi dividido em fases. As gravações na zona rural do município de Rio de Contas constituíram a fase 1 (1996-2000). Essa microrregião fez parte do ciclo de mineração na ocupação do

³ Das frentes de trabalho expostas, a *d* é a que menos tem sido realizada até o momento.

território baiano. Os povoados escolhidos para a gravação das amostras nesse município foram as comunidades de Mato Grosso e de Barra/Bananal. Na fase 2, foram gravadas entrevistas em Jeremoabo, município importante no processo de urbanização do interior do estado nos chamados “ciclos” da agropecuária (séculos XVII a XIX). Os povoados escolhidos para a gravação das amostras nesse segundo município foram Tapera, Lagoa do Inácio e Casinhas, por comporem uma área passível de haver diversidade étnica devido à formação étnico-histórica e social dessas comunidades. Há indícios de que a população seja formada por predominância de negros em Casinhas, por descendentes de índios em Tapera e por mestiços brancos em Lagoa do Inácio, o que poderia favorecer a existência de variações linguísticas (ALMEIDA; CARNEIRO, 2008). Assim, as comunidades das fases 1 e 2 contam com formação diferenciada, a saber, em Rio de Contas: Barra/Bananal (formadas por negros) e Mato Grosso (formada por brancos); e em Jeremoabo: Tapera (formada por remanescentes de índios), Casinhas (remanescentes de quilombolas) e Lagoa do Inácio (formada por mestiços e brancos).

Já na fase 3, as atenções voltaram-se para a comunidade urbana de Feira de Santana, município de maior representatividade demográfica e econômica da imensa faixa do semiárido baiano. Foi constituída amostra com um acervo de 72 entrevistas (quatro informantes por faixa etária, dois homens e duas mulheres) que foram gravadas com o propósito de ser representativas da comunidade de fala dessa cidade.⁴ Para tanto, foi imprescindível o embasamento em estudos sobre a sócio-história do município, os quais assinalaram, por exemplo, a necessidade de analisar a relação entre migração e caracterização sociolinguística na língua falada nessa localidade. Assim, as fases do referido projeto são apresentadas de forma resumida:

⁴ Mais informações podem ser obtidas em Araujo e Almeida (2014).

- a) Fase 1 (1996-2000): amostras de fala das microrregiões da Chapada Diamantina e Piemonte da Diamantina. Foram realizadas gravações nas zonas rurais dos municípios de Caem e Rio de Contas. Zona de mineração.
- b) Fase 2 (2000-2002): amostras de fala das comunidades rurais da região nordeste do estado da Bahia, zona de agropecuária. Gravaram-se inquéritos nos seguintes povoados pertencentes ao município de Jeremoabo: Lagoa do Inácio, Tapera e Casinhas.
- c) Fase 3 (2008-2014): amostras de falas urbana e rural de Feira de Feira de Santana, sendo 12 da fala rural (município da Matinha) e 60 da fala urbana (48 da norma popular e 12 da fala culta).⁵

Frutos dessas fases do projeto, já foram realizados diversos trabalhos científicos, a exemplo de monografias de final de curso, dissertações de mestrado, teses de doutorado, artigos, capítulos e livros. São alguns exemplos desses trabalhos:

– Dissertação de mestrado de Neila Maria Oliveira Santana, intitulada *A indeterminação do sujeito no português rural do semiárido baiano*, Universidade Federal da Bahia (UFBA), 2006.

– Dissertação de mestrado de Juvanete Ferreira Alves Brito, intitulada *O objeto direto (ana)fórico no falar rural baiano: um estudo sociolinguístico*, Universidade Federal de Minas Gerais (UFMG), 2010.

– Dissertação de mestrado de Hilmara Moura de Jesus, intitulada *A regência variável do verbo ir de movimento em comunidades rurais do semiárido baiano*, UFBA, 2011.

⁵ A fase 4 do projeto, cuja ênfase seria em comunidades indígenas, foi prevista, mas ainda não foi executada.

Diálogos Linguísticos e Filológicos

variação, cultura escrita e léxico

– Dissertação de mestrado de Evani Rodrigues, intitulada *Concordância nominal de gênero (em comparação com a fala luan-dense)*, UEFS, 2012.

– Dissertação de mestrado de Adriana de Santana Soares, intitulada *As pretônicas médias em comunidades rurais do semiárido baiano*, UFBA, 2004.

– Livro *Variação linguística no semiárido baiano*, pela UEFS Editora, 2014.

– Tese de doutorado de Silvana Silva de Farias Araujo, intitulada *A concordância verbal no português falado em Feira de Santana-Ba: sociolinguística e sócio-história do português brasileiro*, UFBA, 2014.

– Livro *Variação linguística em Feira de Santana*, pela UEFS Editora, 2016.

Esses estudos, sem dúvida, contribuem para desfazer muitos preconceitos em torno da face popular do PB e trazem valiosas contribuições para o desvendamento da identidade linguística brasileira, discutindo, por exemplo, se ainda há uma bipolarização de normas ou se está ocorrendo uma aproximação entre as duas principais normas sociolinguísticas do PB.

O Projeto *Corpus* Eletrônico de Documentos Históricos do Sertão (CE-DOHS)⁶ tem por objetivo realizar a edição digital de textos do Banco de Documentos Históricos do Sertão (DOHS), do Projeto Vozes do Sertão em Dados: História, Povos e Formação

⁶ O Projeto CE-DOHS é financiado pela Fundação de Amparo à Pesquisa do Estado da Bahia (Fapesb), sob o registro 5566-2010 (UEFS). É coordenado por Zenaide de Oliveira Novais Carneiro e Mariana Fagundes de Oliveira Lacerda e pode ser acessado no endereço <www.uefs.br/cedohs> (cf. UNIVERSIDADE ESTADUAL DE FEIRA DE SANTA, 2011).

do Português Brasileiro,⁷ bem como sua anotação morfológica e sintática, elaborando um *corpus* diacrônico anotado que sirva como recurso eletrônico para o estudo linguístico do português brasileiro. Trata-se, sobretudo, de documentação epistolar manuscrita dos séculos XIX e XX, representativa tanto do português popular brasileiro como do português culto brasileiro; há também livros manuscritos, além de textos impressos e orais. Hoje, contando com melhores recursos tecnológicos no universo das humanidades digitais, os bancos de textos disponibilizam não somente edições semidiplomáticas, em *portable document format* (PDF), mas também edições digitais – a partir do estabelecimento de redes entre projetos que desenvolvem a Linguística de *Corpus* e a Linguística Computacional –, que servem como recurso eletrônico para estudos linguísticos, entre outros. Como se vê,

Do feliz conagraçamento entre as mais recentes tecnologias e a antiga Filologia, surgiu um novo universo de possibilidades para a preservação, disponibilização e análise de textos antigos, universo em que é possível oferecer ao leitor mais de uma edição do mesmo texto, permitindo que tenha ao seu dispor o texto editado, em diferentes versões, e o seu original. (GONÇALVES; BANZA, 2013, p. 4).

Como resultado da primeira fase de pesquisa, o Projeto CE-DOHS já disponibiliza diversos acervos, sobretudo de cartas manuscritas, organizando-as por grau de escolaridade e de habilidade com a escrita (MARQUILHAS, 2000). São 1.084 cartas partculares (1808-2000), num total de 350.850 palavras, escritas por 422

⁷ O Projeto Vozes do Sertão em Dados: História, Povos e Formação do Português Brasileiro (Conselho Nacional de Desenvolvimento Científico e Tecnológico – CNPq – processo 401433/2009-9; Conselho de Ensino, Pesquisa e Extensão – ConsEPE – 102/2009) é coordenado por Zenaide de Oliveira Novais Carneiro.

remetentes (nascidos entre 1724 e 1980), extraída a maior parte de Carneiro (2011). O acervo de cartas ainda está sendo ampliado.⁸

Recentemente foi iniciada a edição, a cargo de Priscila Tuy, do Acervo da Família Estrela Tuy, composto por 106 cartas pessoais inéditas, escritas durante século XX, entre as décadas de 1930 e 1980, por remetentes baianos de origem interiorana e rural e, em sua maioria, com baixa escolaridade. As cartas são em sua totalidade oriundas da Bahia, de pequenos municípios e fazendas pertencentes às imediações desses municípios, além de cartas providas de Salvador. O acervo está organizado da seguinte maneira:

– Cartas de Amor: trata-se de 31 cartas trocadas entre Antonio Carneiro da Silva Tuy e Maria de Souza Estrela (Maria Estrela Tuy). Dessas, 26 foram trocadas durante seu noivado, escritas entre 1952 e 1953, e cinco cartas foram trocadas entre 1956 e 1968, depois de casados.

– Cartas para Antonio e Maria: são 53 cartas enviadas a Antonio Carneiro da Silva Tuy e Maria Estrela Tuy por familiares e amigos.

– Cartas para outros destinatários: são 22 cartas e rascunhos de cartas escritas por Antonio Carneiro da Silva Tuy, Maria Estrela Tuy e outros familiares para amigos e demais entes da família. As cartas foram depositadas ao longo dos anos em uma arca construída pelo próprio Antonio Carneiro da Silva Tuy, durante sua juventude, na Fazenda Bom Jardim, por volta de 1940.

Na segunda fase do projeto, que está em andamento, o número de documentos tem sido ampliado, tanto manuscritos como impressos, com inserção, ainda, de amostras de fala, organizadas no banco por comunidade, tipo de contato linguístico e vertente (popular e “cult”). Essa ampliação do *corpus* favorece essencialmente uma

⁸ Mais informações sobre o Projeto CE-DOHS e sobre a metodologia adotada em seu âmbito podem ser obtidas em Lacerda, Carneiro e Santiago (2016, p. 127-143).

linguística descritiva, fortemente apoiada pelas novas tecnologias, e permite tomar como ponto de partida da descrição a análise de quantidade significativa de dados autênticos, à semelhança do que se faz noutros domínios científicos.

Todo o material do CE-DOHS – representativo de variedades diacrônicas do PB, de diferentes regiões do país e de graus de escolaridade distintos – está sendo preparado para a anotação morfossintática, que manterá a maioria das características do padrão de anotação existente e permitirá a busca automática de dados, o que facilitará o estudo linguístico dos acervos, principal objetivo do projeto. O material disponível no banco atende, entretanto, não somente a pesquisadores interessados em análises de aspectos linguísticos, mas também a interessados em aspectos da difusão de escrita, leitura, transmissões textuais; e aspectos históricos, políticos e econômico-sociais, entre outros. Em 98% dos acervos, é possível determinar *onde, quando, por quem*. No projeto CE-DOHS, os textos-fonte são apresentados em edição semidiplomática, segundo as normas de transcrição do Projeto para a História do Português Brasileiro (2010) (PHPB), sendo oferecidas informações sobre os documentos, sua descrição extrínseca e intrínseca e, sempre que possível, dados biográficos sobre os autores ou, no caso das cartas, os remetentes e os destinatários, como nome, origem, idade, nível de escolaridade, profissão, estado civil, etc.

A codificação dos dados, textuais e extratextuais (ou metadados), é feita com o uso da ferramenta eDictor, o que possibilita a conversão dos textos para diferentes formatos (como .txt, .xml e .html) e evita problemas de processamento eletrônico. As edições filológicas, fidedignas ao texto original, realizadas segundo critérios de transcrição bem definidos, ganham, nos *corpora* digitais, uma versão modernizada, com a padronização da grafia e da acentuação e o desenvolvimento de abreviaturas – todas as alterações ficam

visíveis para o leitor, o que possibilita o controle e mapeamento das intervenções realizadas nos textos, garantindo a recuperabilidade das formas originais. Respeitam-se, entretanto, na edição digital, as mudanças de parágrafo, de linha, as correções do autor, os acidentes do suporte, a orientação da escrita, etc. Com isso, oferece-se uma versão eletrônica de textos sem perder o rigor filológico.⁹

Até o presente, a equipe de pesquisadores do CE-DOHS ocupou-se com a edição filológica e modernizada dos textos (sobretudo com esta, já que a maior parte dos textos que integram o projeto já se encontravam filologicamente editados no Banco DOHS, no âmbito do Projeto Vozes do Sertão em Dados), o que demandou bastante tempo.

A revisão das edições modernizadas – 1.037 cartas manuscritas, mais livros manuscritos, textos impressos e amostras de fala – é feita manualmente, exigindo bastante dedicação da equipe. Todos os textos do banco já passaram pela primeira revisão. Atualmente, passam pela revisão final, antes de receberem a anotação linguística. O CE-DOHS, no processo de anotação linguística dos textos, seguirá os mesmos padrões utilizados por outros projetos de *corpora* eletrônicos, a exemplo do *Corpus* Histórico do Português Tycho Brahe (UNIVERSIDADE ESTADUAL DE CAMPINAS, [20--?]), que é o maior *corpus* eletrônico anotado de textos históricos em português.

Desse modo, as etapas básicas que as equipes de pesquisadores de projetos de *corpora* anotados seguem são essas: anotação de edição, anotação morfossintática e anotação sintática. Na primeira etapa, é utilizado o eDictor; trata-se da codificação de informações

⁹ Por ocasião do *workshop* Construction and Use of Large Annotated Corpora, realizado na Universidade Estadual de Campinas (Unicamp) em 2013, pela equipe do Projeto *Corpus* Histórico do Português Tycho Brahe, do qual pesquisadores de diversos projetos de *corpora* eletrônicos participaram – entre eles, o CE-DOHS –, reafirmou-se a importância de esses projetos seguirem os mesmos padrões de edição digital e de anotação morfossintática, tendo em vista a praticidade do trabalho e a integração dos pesquisadores.

sobre o texto original, sobre decisões editoriais ou sobre a estrutura do texto. Essa etapa, semiautomática, já foi vencida pelo CE-DOHS. A próxima etapa, automatizada, é a de anotação morfossintática, com uso do programa desenvolvido por Faria, Kepler e Paixão de Sousa (2007; 2010), um analisador morfossintático automático, com taxa de acerto de 95%, acoplado ao eDictor. Os erros possíveis de etiquetagem, como já dito, devem ser corrigidos manualmente pelo linguista (o código de etiquetas do eDictor baseia-se no sistema de anotação manual dos Penn Corpora of Historical English (KROCH; SANTORINI; DIERTANI, 2010), da Universidade da Pensilvânia, Estados Unidos. Esse sistema, para adequar-se às peculiaridades da gramática do português, sofreu pequenas alterações.

A anotação sintática, também automatizada, é a terceira e última etapa na constituição de *corpora* anotados; ela diz respeito à identificação e à codificação da estrutura sintagmática do texto. É uma tarefa complexa, mais do que a etiquetagem morfossintática, e exige um *parser* ou, na forma aportuguesada, um parseador, que realiza a análise sintática, reconhecendo identidades em sequências linearmente dispostas e padrões de agrupamentos hierárquicos. Para alcançar seu objetivo – que é a busca automática de dados para estudo da história do PB – a equipe de pesquisadores do CE-DOHS deve cumprir essas duas últimas etapas: a anotação morfossintática e a anotação sintática do *corpus*. O trabalho, até o presente, foi imenso – e ainda há muito a ser feito.

O projeto de pesquisa *Em busca das raízes do português brasileiro* (TEIXEIRA, 2009), idealizado e coordenado pela professora Eliana Pitombo Teixeira no período de 2008 a 2016 e atualmente coordenado pela professora Silvana Silva de Farias Araujo, tem por objetivo consolidar na UEFS a pesquisa linguística, sobretudo no que se refere à história do PB. São os seguintes os seus objetivos específicos: a) formar banco de dados de falantes nativos e

não-nativos do português da comunidade de Luanda; b) descrever e analisar o material coletado nos níveis lexical e morfossintático; c) comparar os resultados dessas descrições e análises com as do PB; e d) contribuir para reflexões em torno da formação sócio-histórica do português do Brasil.

Como bem colocou Mattos e Silva (2009, p. 11) em sua pauta de estudos sobre a história do português popular brasileiro, “o caminho será mais espinhoso, mas não impossível” Ainda que espinhosamente, a África tornou possível o cumprimento dessa pauta. Caminho “espinhoso”, sim, porque Luanda é uma cidade imensa e desordenada, cujo custo de vida é o mais alto do mundo para estrangeiros. Os propósitos, contudo, valeram qualquer sacrifício.

Inserida no grupo de pesquisa Constituição, Variação e Mudança do/no Português Brasileiro”, vinculado ao Nelp, essa investigação é tanto necessária quanto oportuna. Necessária porque contribui para ampliar o debate sobre a gênese da variedade vernacular do português do Brasil e oportuna na medida em que vai ao encontro da política educacional do governo, que, ao editar a Lei nº 10.639, de 9 de janeiro de 2003, promove a valorização e a visibilidade nos espaços escolares/acadêmicos da contribuição africana à cultura brasileira.

A cidade de Luanda, fundada em 1575 pelo português Paulo Dias de Novaes, era simplesmente uma feitoria, criada com um único objetivo: fornecer escravos para lavouras e engenhos do Brasil. Era lá que eram “armazenados” os escravos capturados em toda a região, em barracões próximos ao porto, no qual ficavam à espera de navios que os transportariam para o Novo Mundo.

Inicialmente, de 1500 a 1640, quem fazia o tráfico eram portugueses e tomistas (pessoas nascidas em São Tomé). Além do porto de Luanda, a escravaria saía também dos portos de Cabinda, Benguela, Cuanza (em Angola), Congo e Loango, esses dois últimos ao norte

de Luanda. Os escravos destinavam-se à Costa do Ouro, a São Tomé, Lisboa, Sevilha, América Espanhola, Bahia e Pernambuco. Entre 1641 e 1648, portanto durante sete anos, Angola foi ocupada pelos holandeses, que tomaram o comércio de escravos para as suas possessões na Bahia e em Pernambuco. De 1650 a 1700 o tráfico foi feito exclusivamente por baianos e pernambucanos, os destinos eram o Nordeste e o Sudeste. Até o final do tráfico, em 1850, o comércio de escravos ficou nas mãos de brasileiros (fluminenses, pernambucanos e baianos).¹⁰

O quimbundo serviu de língua franca tanto entre os comerciantes de escravos e os *sobas* (chefes políticos locais) quanto entre os escravos que falavam outras línguas. A captura de escravos era feita basicamente no interior, de onde vinham em longa caminhada para Luanda ou Benguela. Nessas cidades portuárias, passavam meses à espera dos navios, enfrentando, depois, uma longa e penosa viagem para a América; portanto, precisavam de uma língua para se comunicar. Esta foi o quimbundo, língua da colonização e do tráfico.

Voltemos para o lado de cá do Atlântico. Chegados ao Brasil, os escravizados continuaram a usar o quimbundo entre si. Uma evidência disso é a *Arte da língua de Angola*, uma gramática do quimbundo escrita na Bahia pelo jesuíta Pedro Dias, a partir da colaboração de negreiros da Bahia, africanos ladinos e missionários, exatamente no século XVII.¹¹ Ao aprenderem o português no Brasil já adultos e sem nenhuma instrução formal, naturalmente, traços de suas gramáticas foram transferidos para a língua do colonizador. Esse foi o modelo a que tiveram acesso os descendentes dos escravos e os novos escravos que chegavam ao Brasil, o que remete à hipótese de transmissão linguística irregular (LUCCHESI, 2009).

¹⁰ Cf. TRÁFICO TRANSATLÂNTICO DE ESCRAVOS, 2008.

¹¹ Cf. ALENCASTRO, 2009, p. 22.

Em que pese o fato de nem todos os escravizados chegados ao Brasil terem sido naturais de Angola, no século XVII e na primeira metade do XVIII predominaram os angolanos. Além disso, até o final do tráfico, o Brasil recebeu escravizados provenientes de Angola, como afirma Pessoa de Castro (2005). Em suma, queremos sublinhar dois fatos muito pertinentes: a anterioridade e o grande fluxo dos escravizados angolanos naquele século, já que foi nesse período que a variedade brasileira do português começou a se formar e a se expandir. Como evidência, a palavra do padre Antônio Vieira (apud ARAUJO, 1999, p. 254): “quando nessa grande comunidade é já tão pouco geral a chamada língua geral do Brasil que são mui contados aqueles em que se acha?”. Daí o protagonismo do quimbundo na constituição do português popular do Brasil.

Por tudo que foi colocado anteriormente é que se nomeia o projeto: “Em busca das raízes do português brasileiro” – pois a ideia é observar o que resultou do contato dessas línguas com o português europeu e comparar com o PB. Dentre os estudos já realizados com o *corpus*, destacamos alguns, que listamos por ordem cronológica:¹²

– “O pronome *você* no português de Luanda”, capítulo de Eliana Pitombo Teixeira no livro *A língua portuguesa no mundo*, organizado por Maria Célia Lima Fernandes e colaboradores em 2011;

– “A indeterminação do sujeito no português de Luanda: uma comparação com o português do Brasil no século XIX”, artigo de Eliana Pitombo Teixeira e Norma Lucia Fernandes Almeida publicado na revista *Papia* em 2011;

– “O objeto nulo no português falado em Angola”, trabalho da disciplina Variedades do Português apresentado por Huda Santiago e Jadione Cordeiro em 2011;

¹² Alguns desses estudos estão reunidos na coletânea *Diálogos entre África e Brasil: o português daquém e d'além-mar*, publicada pela UEFS e organizada por Eliana Pitombo Teixeira e Silvana Silva de Farias Araujo (2017).

– “A concordância nominal de gênero no português: uma comparação com o português de Luanda”, dissertação de mestrado em Estudos Linguísticos de Evani Rodrigues, apresentada na UEFS em 2012;

– “Aspectos da pronominalização no português falado em Luanda”, capítulo de Eliana Pitombo Teixeira presente no livro *Sociolinguística: estudo da variação, da mudança e da sócio-história do português brasileiro*, organizado por Norma da Silva e colaboradores em 2013;

– “Português brasileiro e português angolano: a concordância nominal de número”, dissertação de mestrado em Estudos Linguísticos de Dayane Lemos, apresentada na UEFS em 2014;

– “A posição dos clíticos no português falado em Luanda”, comunicação de Norma Lucia Fernandes e Eliana Pitombo Teixeira, apresentada no VIII Encontro da Associação Brasileira de Estudos Crioulos e Similares (Abecs) em 2014;

– “Amanhã vais na panela: um estudo sobre a regência do verbo ir de movimento no português de Luanda”, dissertação de mestrado em Estudos Linguísticos de Maria Rosane Passos, apresentada na UEFS em 2015.

Os resultados de todos esses estudos têm apresentado convergências com o português vernacular brasileiro e levado a descobertas interessantes. Por exemplo, Teixeira (2008), em seu primeiro trabalho com os dados de Luanda, estudou o uso pronome *você* substituindo o *tu* e observou que são as pessoas da base da pirâmide social que o usam com mais frequência. Em estudo diacrônico sobre a mudança *tu* → *você* no Brasil, no século XIX e na primeira metade do XX (cf. TEIXEIRA, 2002), mostra-se que os usuários desse pronome eram também pessoas da classe baixa. Diante disso, a autora chegou à conclusão de que essas pessoas usavam tal forma de tratamento porque

eram assim tratados pelos colonizadores, tanto no Brasil como em Angola, ou seja, essa era a forma a que eles tiveram acesso e que, com o tempo, foi se expandido para outros estratos sociais.

Em suma, o que se observa em relação ao efeito do contato do português europeu com as línguas africanas em Angola pode espelhar o que aconteceu no Brasil em tempos pretéritos.

CONSIDERAÇÕES FINAIS

A temática focalizada neste texto traz um vasto campo para pesquisas. É tarefa para muitos estudiosos, munidos de diferentes arsenais teórico-metodológicos e com diferentes fontes de dados. Somente assim poderá a constituição do PB ser mais bem avaliada. Considerando os aspectos aqui discutidos, entende-se que o debate sobre a história sociolinguística do PB é fundamental para um entendimento contextualizado acerca da identidade linguística brasileira atual. Assim, evitam-se afirmações genéricas e focalizam-se os fenômenos linguísticos encaixados na estrutura histórico-social mais ampla da comunidade de fala analisada. Considera-se ainda que a hipótese interpretativa da *transmissão linguística irregular* é consistente para a formação do PB. Dessa forma, a constituição de *corpora* é um excelente e imperioso meio para melhor entender as complexas questões da formação e da caracterização atual do PB.

REFERÊNCIAS

ALENCASTRO, Luiz Felipe de. Os africanos e as falas africanas no Brasil. In: GALVES, Charlotte; GARMES, Hélder; RIBEIRO, Fernando Rosa (Org.). *África-Brasil: os caminhos da língua portuguesa*. Campinas: Editora da Unicamp, 2009.

ALMEIDA, Norma Lúcia Fernandes de; CARNEIRO, Zenaide de Oliveira Novais (Org.). *Coleção amostra da língua falada no semiárido baiano*. Feira de Santana: UEFS Editora; Salvador: Fapesb, 2008.

ARAUJO, Jorge. *Profecias morenas: discurso do eu e da pátria em Antonio Vieira*. Salvador: Academia de Letras da Bahia, 1999.

ARAUJO, Silvana Silva de Farias. *A concordância verbal no português falado em Feira de Santana-BA: sociolinguística e sócio-história do português brasileiro*. 2014. 342 f. Tese (Doutorado em Língua e Cultura) – Instituto de Letras, Universidade Federal da Bahia, Salvador, 2014.

ARAUJO, Silvana Silva de Farias. O português popular do semiárido baiano: fundamentos teóricos, sócio-históricos e empíricos. In: LOPES, Norma da Silva; OLIVEIRA, Josane Moreira; PARCERO, Lúcia Maria de Jesus (Org.). *Estudos sobre o Português do Nordeste: língua, lugar e sociedade*. São Paulo: Blucher, 2017.

ARAUJO, Silvana Silva de Farias; ALMEIDA, Norma Lucia Fernandes de. O projeto A Língua Portuguesa no Semiárido Baiano – Fase 3: critérios de constituição e da amostragem do banco de dados. In: FREITAG, Raquel Meister Ko (Org.). *Metodologia de coleta e manipulação de dados em Sociolinguística*. São Paulo: Edgard Blücher, 2014. p. 27-47.

BRASIL. Ministério da Educação. Lei nº 10.639, de 9 de janeiro de 2003. Altera a Lei nº 9.394, de 20 de dezembro de 1996, que estabelece as diretrizes e bases da educação nacional, para incluir no currículo oficial da rede de ensino a obrigatoriedade da temática “História e Cultura Afro-Brasileira”, e dá outras providências.

Diálogos Linguísticos e Filológicos

variação, cultura escrita e léxico

Diário Oficial da União, Brasília, 10 jan. 2003. Disponível em:
<<http://bit.ly/2z0UkRl>>. Acesso em: 28 jun. 2018.

CARNEIRO, Zenaide de Oliveira Novais. (Org.). *Cartas brasileiras (1809-2000)*: coletânea de fontes para o estudo do português. Feira de Santana: UEFS Editora, 2011.

CARNEIRO, Zenaide de Oliveira Novais; LACERDA, Mariana Fagundes de Oliveira. *Publica-se em Feira de Santana*: das cartas de leitores e redatores e dos anúncios em *O Progresso e Na Folha do Norte* (1901-2006). Feira de Santana: UEFS Editora, 2012.

GONÇALVES, Maria Filomena; BANZA, Ana Paula. Fontes metalinguísticas para a história do português clássico. In: GONÇALVES, Maria Filomena; BANZA, Ana Paula. *Patrimônio textual e humanidades digitais*: da antiga à nova filologia. Évora: Cidehus, 2013. p. 3-9.

HOUAISS, Antônio. *O português do Brasil*. Rio de Janeiro: Unibrade; Centro de Cultura, 1985.

KROCH, Anthony; SANTORINI, Beatrice; DIERTANI, Ariel. The Penn Parsed Corpus of Modern British English (PPCMBE2). In: ANNUAL MEETING OF THE ASSOCIATION FOR COMPUTATIONAL LINGUISTICS, 52., 23-25 jun. 2014, Baltimore. *Proceedings...* Filadélfia: University of Pennsylvania, 2016. CD-ROM. Disponível em: <<https://goo.gl/XoMmNB>>. Acesso em: 13 jul. 2018.

LABOV, William. *Padrões sociolinguísticos*. Tradução Marcos Bagno, Maria Marta Pereira Scherre, Caroline Rodrigues Cardoso. São Paulo: Parábola, 2008.

LACERDA, Mariana Fagundes de Oliveira; CARNEIRO, Zenaide de Oliveira Novais; SANTIAGO, Huda da Silva. Corpus eletrônico de documentos históricos do sertão: as cartas de inábeis. *A cor das Letras*, Feira de Santana, v. 17, p. 127-143, 2016.

LOBO, Tânia Freire Conceição. *Para uma sociolinguística histórica do português do Brasil*: edição filológica e análise linguística de cartas particulares do Recôncavo da Bahia, século XIX. Tese (Doutorado em Filologia e Língua Portuguesa) – Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo, 2001.

LUCCHESI, Dante. As duas grandes vertentes da história sociolinguística do Brasil (1500-2000). *Delta*, São Paulo. v. 17, n. 1, p. 97-132, 2001.

LUCCHESI, Dante. História do contato entre línguas no Brasil. In: Lucchesi, Dante; Baxter, Alan; Ribeiro, Ilza (Org.). *O português afro-brasileiro*. Salvador: Edufba, 2009. p. 41-73.

LUCCHESI, Dante. *Língua e sociedade partidas*: a polarização sociolinguística no Brasil. São Paulo: Contexto, 2015.

LUCCHESI, Dante. Variação e norma: elementos para uma caracterização sociolinguística do português do Brasil. *Revista Internacional de Língua Portuguesa*, Lisboa, n. 12, p. 17-28, 1994.

MARQUILHAS, Rita. *A faculdade das letras*: leitura e escrita em Portugal no séc. XVII. Lisboa: Imprensa Nacional; Casa da Moeda, 2000.

MATTOS E SILVA, Rosa Virgínia. Da sócio-história do português brasileiro para o ensino do português hoje. In: Azeredo, José Carlos

de (Org.). *Língua portuguesa em debate*. Conhecimento e ensino. Petrópolis: Vozes, 2000. p. 19-33.

MATTOS E SILVA, Rosa Virgínia. De fontes sócio-históricas para a história social linguística do Brasil: em busca de indícios. In: MATTOS E SILVA, Rosa Virgínia (Org.). *Para a história do português brasileiro: primeiros estudos*. São Paulo: Humanitas; Fapesp, 2001. v. 2. t. 1-2. p. 275-302.

MATTOS E SILVA, Rosa Virgínia. Variação, mudança e norma (movimentos no interior do português brasileiro). In: BAGNO, Marcos (Org.). *Linguística da norma*. São Paulo: Loyola, 2002, p. 291-316.

MATTOS E SILVA, Rosa Virgínia. A generalizada difusão da língua portuguesa no território brasileiro. In: MATTOS E SILVA, Rosa Virgínia. *Ensaio para uma sócio-história do português brasileiro*. São Paulo: Parábola, 2004, p. 91-108.

MATTOS E SILVA, Rosa Virgínia. *Caminhos da linguística histórica: ouvir o inaudível*. São Paulo: Parábola, 2008.

MATTOS E SILVA, Rosa Virgínia. Para a história do português culto e popular brasileiro: sugestão para uma pauta de pesquisa. *Cadernos de Letras da UFF*, Niterói, v. 34, p. 11-30, 2009.

MUSSA, Alberto B. N. *O papel das línguas africanas na história do português do Brasil*. 1991. Dissertação (Mestrado em Língua Portuguesa) – Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 1991.

PAGOTTO, Emílio. Norma e condescendência: ciência e pureza. *Línguas e instrumentos linguísticos*, São Paulo, n. 2, p. 49-68, 1998.

PAIXÃO DE FARIA, Pablo Picasso Feliciano de; KEPLER, Fabio Natanael; PAIXÃO DE SOUSA, Maria Clara. E-dictor: novas perspectivas na codificação e edição de corpora de textos históricos. In: ENCONTRO DE LINGUÍSTICA DE CORPUS, 8., 2009, Rio de Janeiro. *Anais...* Rio de Janeiro: UERJ, 2009. p. 69-105.

PAIXÃO DE FARIA, Pablo Picasso Feliciano de; KEPLER, Fabio Natanael; SOUSA, Maria Clara Paixão de. *An integrated tool for annotating historical corpora*. In: LINGUISTIC ANNOTATION WORKSHOP, 4.; ANNUAL MEETING OF THE ACL, 48., 2010, Uppsala. *Proceedings...* Stroudsburg: ACL, 2010. p. 217-221.

PESSOA DE CASTRO, Ieda. *Falares africanos na Bahia: um vocabulário afro-brasileiro*. 2. ed. Rio de Janeiro: Topbooks, 2005.

PROJETO PARA A HISTÓRIA DO PORTUGUÊS BRASILEIRO. *Plataforma de corpora do PHPB*. [S. l.], 2010. Disponível em: <<http://bit.ly/2tIXKBZ>>. Acesso em: 3 maio 2017.

TEIXEIRA, Eliana Pitombo. *Era uma vez você*. 2002. Tese (Doutorado em Letras e Linguística) – Instituto de Letras, Universidade Federal da Bahia, Salvador, 2002.

TEIXEIRA, Eliana Pitombo. *Em busca das raízes do português brasileiro*. Projeto de pesquisa. Feira de Santana: Universidade Estadual de Feira de Santana, 2009.

TEIXEIRA, Eliana Pitombo. O pronome você no português de Luanda. In: LIMA-HERNANDES, Maria Célia et al. (Org.). *A língua portuguesa no mundo*. São Paulo: FFLCH-USP, 2008.

TEIXEIRA, Eliana Pitombo; ARAUJO, Silvana Silva de Farias (Org.). *Diálogos entre Brasil e Angola: o português d'aquém e d'além mar*. Feira de Santana: UEFS, 2017.

Diálogos Linguísticos e Filológicos
variação, cultura escrita e léxico

TRÁFICO TRANSATLÂNTICO DE ESCRAVOS. *Banco de dados de viagens*. [S. l.], 2008. Disponível em: <<http://bit.ly/2KvrKvy>>. Acesso em: 28 jun. 2018.

UNIVERSIDADE ESTADUAL DE CAMPINAS. Instituto de Estudos da Linguagem. *Corpus Histórico do Português Tycho Brahe*. Campinas, [20--?]. Disponível em: <<http://bit.ly/2yOfcLb>>. Acesso em: 10 ago. 2016.

UNIVERSIDADE ESTADUAL DE FEIRA DE SANTANA. *Corpus eletrônico de documentos históricos do sertão*. Coordenado por Zenaide de Oliveira Novais Carneiro e Mariana Fagundes de Oliveira. Feira de Santana, 2011. Disponível em: <<http://bit.ly/2lBH8IP>>. Acesso em: 10 ago. 2016.