

O Núcleo de Estudos de Língua Portuguesa da Universidade Estadual de Feira de Santana: centro de linguística histórica no semiárido baiano

The nucleus of Portuguese language studies at the State University of Feira de Santana, a centre of historical linguistics in the semi-arid region of Bahia

Mariana Fagundes de Oliveira Lacerda*

RESUMO

A Linguística Histórica tem encontrado espaço de desenvolvimento no Brasil, em diferentes instituições de Ensino Superior, tendo sido realizadas, no país, três edições do Congresso Internacional de Linguística Histórica (CILH), a primeira delas, em 2009, homenageando a professora Rosa Virgínia Mattos e Silva, na Universidade Federal da Bahia (UFBA), pesquisadora que, ao lado de seus colegas e alunos vinculados ao Programa para a História da Língua Portuguesa (PROHPOR), maior programa de Linguística Histórica do Brasil, criado em 1990, levou a Bahia a ser reconhecida como um Estado forte em estudos da língua portuguesa na perspectiva histórica. E, também no Semiárido baiano, a Linguística Histórica encontrou espaço de desenvolvimento, na Universidade Estadual de Feira de Santana (UEFS), no âmbito do Núcleo de Estudos de Língua Portuguesa (NELP), criado em 1998, cujas pesquisas, até hoje, seguem, fundamentalmente, as proposições de Mattos e Silva, em sua vasta obra. Neste trabalho, procuramos demonstrar, de forma reflexiva, como o NELP/UEFS, que trabalha em parceria com o Projeto Nacional para a História do Português Brasileiro (PHPB), entre outros parceiros, tornou-se um centro de Linguística Histórica, nacionalmente reconhecido, atuando em diferentes frentes,

Recebido em 17 de julho de 2023

Aceito em 19 de dezembro de 2023.

DOI: <https://doi.org/10.18364/rc.2024n67.1389>

* marianafagundes@uefs.br

como a constituição de *corpora* diacrônicos de português brasileiro, o estudo sócio-histórico e o estudo gramatical, realizando trabalhos que se tornaram referência importante e contribuíram para o avanço do conhecimento na área. Trata-se de um texto que apresenta os campos de investigação do NELP e sua metodologia de trabalho, podendo servir a experiência do núcleo, como ele se constitui e empreende suas investigações, como modelo a pesquisadores interessados em pesquisas em Linguística Histórica e História do Português Brasileiro.

PALAVRAS-CHAVE: NELP; Semiárido Baiano; Linguística Histórica.

ABSTRACT

Historical linguistics has found space for its development in various higher education institutions in Brazil. Three editions of the International Congress of Historical Linguistics (CILH) have been held in the country. The first was held in 2009 in honor of Professor Rosa Virgínia Mattos e Silva at the Federal College of Bahia (UFBA), a researcher who, together with her colleagues and students from the Portuguese Language History Program (PROHPOR), founded in 1990 and the largest historical linguistics program in Brazil, has ensured that Bahia is recognized as a strong state in Portuguese linguistics from a historical perspective. And also in the semi-arid region of Bahia, historical linguistics found room to flourish at the State College of Feira de Santana (UEFS) within the Nucleus of Portuguese Language Studies (NELP), founded in 1998, whose research to this day essentially follows the theses of Mattos e Silva in his extensive work. In this paper, we attempt to show how the NELP/UEFS, in collaboration with the National Project for the History of Brazilian Portuguese (PHPB), among others, has become a nationally recognized center for historical linguistics. It works on several fronts, such as the creation of diachronic corpora of Brazilian Portuguese, socio-historical studies and grammatical studies, and carries out work that has become a model and contributes to the advancement of knowledge in the field. This is a text that presents the research areas of the NELP and its working methodology. The experience of the nucleus, how it is constituted and how it conducts its investigations, can serve as a model for researchers interested in research in Historical Linguistics and the History of Brazilian Portuguese.

KEYWORDS: NELP; Bahia semi-arid region; historical linguistics

Palavras iniciais

O I Congresso Internacional de Linguística Histórica (I CILH) foi realizado em 2009, em homenagem à professora Rosa Virgínia Mattos e Silva, cuja obra sobre a história da língua portuguesa, recuando ao período arcaico (XIII-XVI), e sobre a história do português brasileiro (PB) tem reconhecida importância. Foi vinculada ao corpo docente da Universidade Federal da Bahia (UFBA), *campus* de Salvador, capital baiana, que a renomada professora construiu grande parte de sua história acadêmica, em cujas linhas se encontra também a Universidade Estadual de Feira de Santana (UEFS), no sertão da Bahia, com a qual sempre dialogou, instituição que possui, hoje, no seu quadro funcional, no Departamento de Letras e Artes (DLA), ex-alunas da querida pró Rosa, atualmente professoras doutoras, seguindo pelos caminhos da Linguística Histórica...

O Núcleo de Estudos de Língua Portuguesa (NELP) da UEFS – maior núcleo de pesquisa, com estudos na perspectiva histórica, do Nordeste brasileiro – teve sua criação, na década de 90 do século XX, inspirada na obra de Mattos e Silva. Na ocasião do I Ciclo de Palestras do NELP, realizado em julho de 2020, em homenagem à referida professora, tivemos a oportunidade de afirmar que, na UEFS, o maior produto, e mais significativo, que temos, como fruto dos ensinamentos dela, é o NELP.

Neste texto, procuramos demonstrar, de forma reflexiva, como o NELP/UEFS – que trabalha em parceria com o Projeto Nacional para a História do Português Brasileiro (PHPB), entre outros parceiros – tornou-se um centro de Linguística Histórica no Semiárido Baiano, nacionalmente reconhecido, investindo em pesquisas sólidas – que se tornaram modelo e contribuíram para o avanço do conhecimento na área –, com financiamento de diferentes instituições de fomento, considerando, fundamentalmente, as proposições da professora soteropolitana, uma das maiores linguistas brasileiras, em sua vasta obra. Trata-se de um texto que apresenta os campos de investigação do NELP e sua metodologia de trabalho, podendo servir a experiência do

núcleo, como ele se constitui e empreende suas investigações, como modelo a pesquisadores interessados em pesquisas em Linguística Histórica e História do Português Brasileiro.

O texto está organizado nas seguintes partes: seguem estas Palavras Iniciais a seção 1, que apresenta, em síntese, as conquistas do NELP, em 25 anos de história; na seção 2, descrevemos as agendas do núcleo, dando destaque à constituição de *corpus* diacrônico, da qual resulta um banco de dados sociolinguísticos, pioneiro no Nordeste; sobre a exploração do banco de textos do NELP na interface educacional tratamos na seção 3, seguida pelas Palavras Finais e pelas Referências consultadas.

1 O NELP em 25 anos de história

Neste ano, em 2023, o NELP completa 25 anos. E este texto é também uma homenagem ao núcleo, pela sua longevidade e vitalidade.



Figura 1: Logotipo do NELP.

Fonte: <<https://nelp.uefs.br/o-nelp/logotipo/>>

Um núcleo que nasce no interior da Bahia, e vai longe, com uma produção muito significativa, da qual faz parte um banco de dados que se tornou referência importante (conferir seção 2). Na Figura 1, temos o novo logotipo do NELP, criado em 2020, a partir de uma obra do artista Juraci Dórea Falcão, tendo, como inspiração, os sertões baianos, sua principal área de abrangência.

O NELP – inicialmente denominado Núcleo de Estudos de Língua Portuguesa Rural (NELPRU)¹ –, coordenado, desde 2017, pela professora Mariana Fagundes de Oliveira Lacerda, foi criado pelas professoras Zenaide de Oliveira Novais Carneiro e Norma Lucia Fernandes de Almeida, em 1998, no Departamento de Letras e Artes (DLA) da Universidade Estadual de Feira de Santana (UEFS), a partir das pesquisas desenvolvidas no âmbito do projeto A Língua Portuguesa no Semiárido Baiano, por elas coordenado e fruto de estudos iniciados, na instituição, pela professora Ilza Ribeiro e pelo professor Dante Lucchesi, sobre a constituição de banco de dados para investigação da história do português brasileiro.



Figura 2: Linha do tempo do Nelp, card 1.

Fonte: Instagram do NELP <@nucleonelp>

1 Conferir artigo intitulado *O NELPRU (Núcleo de Estudos de Língua Portuguesa Rural): apresentação de alguns resultados*, de Almeida e Carneiro (2003).

Hoje, fazem parte do NELP 14 projetos de pesquisa, coordenados por pesquisadores da UEFS, que trabalham em parceria com outras universidades brasileiras e universidades estrangeiras. O NELP, ao lado de outros núcleos de pesquisa do departamento, fortalece a linha de pesquisa Variação e Mudança, do Programa de Pós-Graduação em Estudos Linguísticos (PPGEL) da UEFS.

Na grande área da Linguística Histórica, o núcleo se dedica, desde sempre, especialmente à história do português brasileiro, com pesquisas desenvolvidas dentro de diferentes quadros teóricos, sobretudo seguindo a tendência de estudo sociologizante, afinando-se com Labov (2008 [1972]), com o qual os valores culturais do falante passam a primeiro plano, fomentando o respeito à diversidade linguística (intimamente relacionada à diversidade sociocultural), em defesa da dignidade humana.

1.1 As agendas de pesquisa e os caminhos teóricos

Parceiro, como já dissemos, do PHPB – criado, em 1997, a partir do Projeto História do Português Paulista (PHPP), coordenado por Ataliba Teixeira de Castilho (CASTILHO, 1998; 2018), e inspirado também no Programa para a História da Língua Portuguesa (PROHPOR), coordenado por Rosa Virgínia Mattos e Silva, na UFBA², com o intuito de “desvendar o passado da língua portuguesa e do português brasileiro (MATTOS E SILVA, 2004, p. 122) – , o NELP segue seu programa investigativo, que busca historiar a formação da identidade linguística do Brasil, por meio dos quatro sistemas que organizam uma língua natural: Gramática, Discurso, Léxico e Semântica. As agendas de trabalho propostas pelo PHPB são: “(1) organização do *corpus* diacrônico; (2) História social do PB; (3) Mudança gramatical do PB; (4) Tradições discursivas: constituição e mudança dos gêneros discursivos. Diacronia dos

2 Hoje, o PHPP é coordenado pelo professor Manoel Mourivaldo Santiago Almeida (<https://phpp.fflch.usp.br/>), e o PROHPOR, pela professora Juliana Soledade Barbosa Coelho (<https://www.prohpor.org/>).

processos constitutivos do texto; (5) História do Léxico” (CASTILHO, 2018, p. 28-29). O núcleo vem investindo esforços, sobretudo, nas três primeiras agendas; no que diz respeito à terceira agenda, os estudos são feitos na perspectiva extrassistêmica (gramaticais e hipergramaticais) e intrassistêmica (de linguística diacrônica).

Podemos dizer, de forma mais detalhada, que o NELP desenvolve suas pesquisas sobre a história social linguística do PB, considerando as proposições de Mattos e Silva, sintetizadas por Lobo (2015, p. 70):

1. A história linguística do Brasil não se restringe à história da língua portuguesa no Brasil, nem à história do português brasileiro.
2. O português brasileiro emerge em contexto multilíngue: o contato linguístico é, pois, elemento constitutivo da sua formação.
3. Na cena linguística do Brasil colonial, destacam-se três atores principais: o português europeu, as línguas gerais indígenas e o português geral brasileiro.
4. Africanos e afrodescendentes foram os principais difusores da língua portuguesa no Brasil e os principais formatadores do português brasileiro em sua variante social majoritária — o português popular brasileiro.
5. O passado sócio-histórico-linguístico do Brasil deverá ser interpretado para a compreensão do português brasileiro «heterogêneo e variável, plural e polarizado» da atualidade. (LOBO, 2015, p. 70)

Mattos e Silva, considerando as vias de pesquisa para “elaborar uma história da língua portuguesa no Brasil”, propostas por Houaiss (1985, p. 31), apresenta quatro campos de pesquisa para investigar “mudanças linguísticas que fizeram e fazem o português brasileiro ter as características que tem, o seu perfil próprio, a sua gramática” (MATTOS E SILVA, 2004, p. 58). Os campos referidos são:

- (a) o campo que se moverá na reconstrução de uma história social linguística do Brasil;
- (b) o campo que se moverá na reconstrução de uma sócio-história linguística ou de uma sociolinguística histórica;
- (c) o campo que se moverá na reconstrução diacrônica no interior das estruturas da língua portuguesa em direção ao português brasileiro;
- (d) o

campo que se moverá no âmbito comparativo entre o português europeu e o português brasileiro (MATTOS E SILVA, 2004, p. 58).

Nesses diferentes campos, o NELP tem pesquisas realizadas. E, para tanto, constituiu, ao longo de 25 anos, desde sua criação, um banco de dados diacrônicos, representativo do PB, com textos manuscritos e impressos, em edição semidiplomática e modernizada (além de amostras de fala da década de 90 do século XX e atuais), como o leitor pode conferir na seção 2. A parceria da Linguística Histórica *Stricto Sensu* com a Filologia é essencial, conforme apresenta Mattos e Silva (2008, p. 10) na figura a seguir:

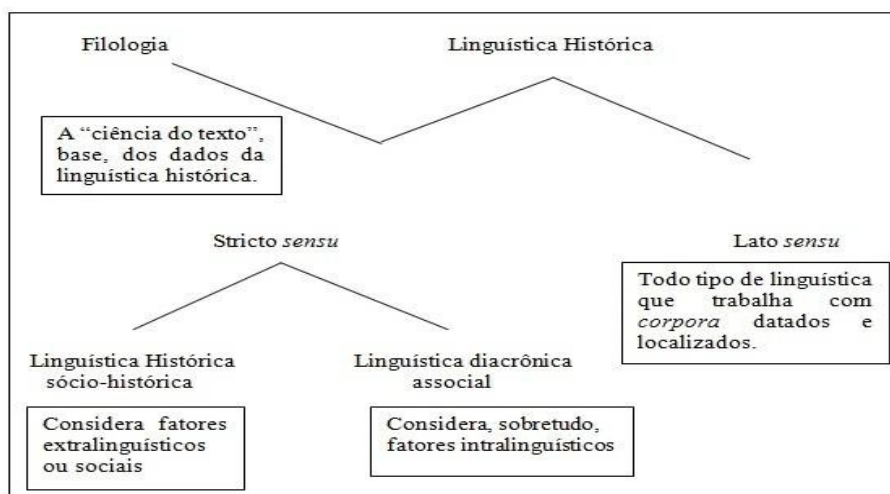


Figura 3: Parceria entre a Linguística Histórica *Stricto Sensu* e a Filologia.

Fonte: Mattos e Silva (2008, p. 10).

É Linguística Histórica *Stricto Sensu* – definida, segundo Mattos e Silva (2008, p. 9), como a vertente que se debruça sobre o que muda nas línguas e como a mudança se dá, dividida em duas orientações, a sócio-histórica (ou perspectiva de estudo extralinguística ou extrassistêmica) e a diacrônica

associal (ou perspectiva de estudo intralinguística ou intrassistêmica) – que o NELP faz, sobretudo, e, portanto, anda de mãos dadas com a Filologia:

Deve-se ressaltar que não se pode nem se deve utilizar qualquer edição de texto do passado para a análise histórico-diacrônica: a edição tem de ser feita com rigor filológico e com objetivo claro de servir a estudos linguísticos; há edições úteis ao historiador ou ao estudioso da literatura ou chamado grande público, mas que, contudo, não devem ser usadas para estudos de história linguística (MATTOS E SILVA, 2008, p. 15).

Dessa forma, fazem parte da equipe do NELP pesquisadores com formação em filologia, e também em Paleografia, dedicados ao trabalho de edição de textos, com o rigor filológico e paleográfico necessário. É nos textos remanescentes do passado, afinal, que encontramos os dados diacrônicos para a pesquisa em Linguística Histórica (MATTOS E SILVA, 1998, p. 106). Como ressalta Paixão de Souza (2006, p. 13), no fazer histórico, seja na história das línguas, seja em qualquer história, não temos acesso direto aos processos abordados, são tempos diferentes, o tempo da análise e o tempo do objeto analisado e “Essa impossibilidade dos métodos de observação imediata deixa duas alternativas principais aos estudos históricos: o recurso à documentação, e o recurso à reconstrução.” O NELP recorre à documentação, estabelecendo uma parceria necessária com a Filologia Textual e a Paleografia de Análise.

Esses dados diacrônicos, como dissemos anteriormente, citando Mattos e Silva, podem ser analisados em diferentes perspectivas, social ou associal, e o NELP segue pelos dois caminhos, ainda que sejam mais frequentes, no seu âmbito, os estudos na tendência sociologizante, tendo grande parte de seus pesquisadores uma concepção social de língua. E a pergunta que fica, à qual Paixão de Souza (2006, p. 27) respondeu bem, é: como teorias com objetos-língua a-históricos ou associais se inserem na grande área da Linguística Histórica?

Podemos entender como mesmo os diferentes quadros teóricos da linguística cujos objetos-língua se constituem como a-históricos realizem suas incursões pela “linguística histórica”: é que, de fato, esses quadros partem de uma perspectiva estritamente cronológica,

ocupando-se primordialmente da sucessão linear de fatos de língua no tempo. (...) Com temporalidade, mas sem historicidade (PAIXÃO DE SOUZA, 2006, p. 27)

Então, como vemos, é possível fazer Linguística Histórica sem considerar a historicidade das línguas ou tomando-a como objeto teórico sem dimensão histórica; escolher esse “lote” do terreno é possível. Como afirma Borges Neto (2004, p. 35-36), esse “loteamento” é um primeiro momento de teorização:

Um erro comum é supor que as divisões da ciência correspondem a divisões naturais da realidade. Isto é equivalente a supor que, pelo fato de alguém ter direitos adquiridos sobre certo território, as fronteiras desse território correspondem a alguma divisão natural. As delimitações dos objetos observacionais não são neutras, ou seja, não é a própria realidade que diz como quer ser seccionada. O “loteamento” do observacional é resultado de um trabalho humano sobre a realidade e, em consequência, já é um primeiro momento de teorização. (BORGES NETO, 2004, p. 35-36)

Tudo depende das escolhas teóricas, que, por sua vez, não são neutras, mas passam pela subjetividade do pesquisador, suas ideologias. O NELP reúne pesquisadores que trabalham, sobretudo, na dimensão do temporal-histórico, fazendo Linguística Histórica, com historicidade, tendo um programa investigativo e proposições bem definidos.

1.2 NELP em números

Nesta seção, apresentaremos números a que o NELP chegou, nesses 25 anos de pesquisas e estudos sobre a histórica social linguística do PB, atuando nas diferentes agendas supracitadas. No quadro seguinte, uma síntese do que consideramos importante registrar:

Categorias	Números
Professores	14
Estudantes de Iniciação Científica	170
Estudantes de Mestrado	62
Estudantes de Doutorado	15
Livros publicados	37
Capítulos publicados	125
Artigos publicados	161
Eventos acadêmicos organizados	16

Quadro 1: Números do NELP em 25 anos de história.³

Fonte: Elaboração própria.

A significativa produção bibliográfica do núcleo, entre livros, capítulos e artigos, deve-se ao intenso trabalho da equipe de professores pesquisadores, em parceria com os estudantes de Graduação e de Pós-Graduação, sem a colaboração dos quais não seria possível fazer tanto. Inclusive os eventos acadêmicos, até o presente promovidos pelo NELP – alguns em parceria com outros núcleos de pesquisa da UEFS e com outras instituições de Ensino Superior e instituições de Ensino Básico –, contaram sempre com a valiosa participação dos estudantes na Comissão Organizadora.

Entre os livros publicados, destacamos, a seguir, a primeira coleção publicada e a mais recente, sobre constituição de *corpus*, agenda de trabalho na qual o NELP é referência nacional e internacional. É a partir do *corpus* constituído que o NELP desenvolve os estudos sócio-históricos e as análises gramaticais.

A coleção mais antiga do núcleo, intitulada *Coleção Amostras da Língua Falada no Semiárido Baiano*, organizada pelas professoras Norma

3 Todas as informações podem ser conferidas no currículo Lattes dos pesquisadores do NELP, cujo link encontra-se disponível no site do núcleo, no menu Equipe/Pesquisadores (<https://nelp.uefs.br/pesquisadores/>).

Lucia Fernandes de Almeida e Zenaide de Oliveira Novais Carneiro, foi publicada em 2008, pela UEFS Editora como resultado do projeto A Língua Portuguesa no Semiárido Baiano, financiado pela Fundação de Amparo à Pesquisa do Estado da Bahia (FAPESB).

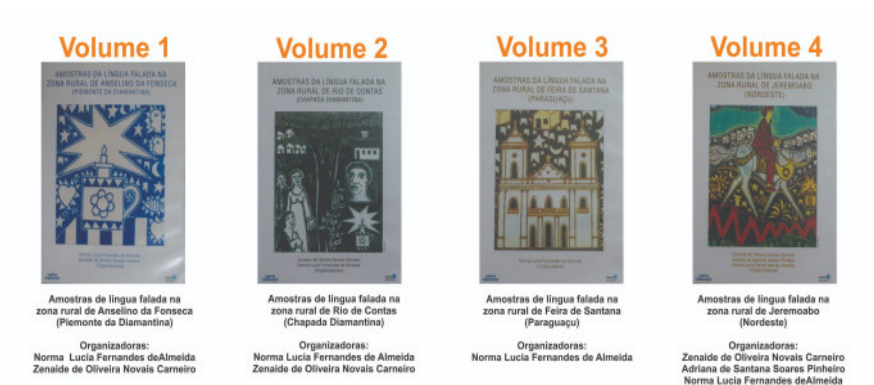


Figura 4: Capas dos quatro volumes da Coleção Amostras da Língua Falada no Semiárido Baiano.

Fonte: <<http://www.uefs.br/cedohs/>>.

Trata-se de coleção com quatro volumes, com inquéritos gravados entre 1997 e 1999, com o objetivo de contribuir para o conhecimento da realidade linguística brasileira e, de forma específica, para o estudo da língua falada em áreas do semiárido baiano. As comunidades foram escolhidas de forma que representassem o avanço da língua portuguesa na Bahia, a partir do século XVII, numa perspectiva sócio-histórica.

As amostras rurais foram compostas em áreas passíveis de diversidade étnica, havendo indícios de que a população fosse formada por remanescentes de quilombo em Casinhas (Jeremoabo); por presença étnica de origem afro-brasileira em Barra/Bananal (Rio de Contas/Região da Chapada Diamantina), Piabas (Caém/Ancelino da Fonseca/Região do Piemonte da Diamantina) e Matinha (Feira de Santana/Região Paraguaçu); por mestiços de brancos de origem portuguesa em Lagoa do Inácio (Jeremoabo) e Mato

Grosso (Rio de Contas). Isto implica a existência de variações linguísticas provenientes de contato linguístico entre línguas distintas, no processo de aquisição de língua (CARNEIRO *et al*, 2016, p. 96).

Por meio dessas gravações com pessoas analfabetas ou de baixo grau de escolarização, em zonas rurais, o NELP oferece à comunidade científica material para estudo do PB, do presente para o passado.

A outra coleção, última publicada pelo núcleo, em 2022, também pela UEFS Editora, é *O sertão por Escrito: edição filológica dos livros do arquivo do Sobrado do Brejo (Bahia)*, organizada pelas professoras Zenaide de Oliveira Novais Carneiro e Mariana Fagundes de Oliveira Lacerda.



Figura 5: Capas dos dois volumes da Coleção O Sertão por Escrito.

Fonte: <<http://www.uefs.br/cedohs/>>

Essa coleção apresenta 2 volumes: o primeiro com as edições fac-similar e semidiplomática do *Livro do Gado*, feitas com auxílio financeiro da FAPESB, no âmbito da dissertação de Mestrado de Elaine Santos (2019),

orientada pela professora Mariana Fagundes de Oliveira Lacerda e defendida no PPGEL da UEFS, e o segundo com as edições fac-similar e semidiplomática do *Livro de Razão*⁴, feitas com auxílio financeiro da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), no âmbito da tese de Doutorado de Adilson Silva (2021), orientado pelas professoras Emília Helena Portella Monteiro de Souza e Zenaide de Oliveira Novais Carneiro e defendida no Programa de Pós-Graduação em Língua e Cultura (PPGLinC) da UFBA.⁵

O NELP tem, nos livros manuscritos do arquivo do Sobrado do Brejo do Campo Seco, os primeiros documentos representativos do período colonial brasileiro a fazerem parte de seu banco de textos.

Documentação rara, de foro privado, esses livros escritos por três gerações, na fazenda de criação do Brejo do Campo Seco, da família Pinheiro Canguçu – povoado de Bom Jesus dos Meiras, Comarca de Rio de Contas (hoje Brumado), sertão baiano – têm inegável valor histórico, econômico, social e linguístico. Como ressalta Algranti (1997, p. 132-133), recuperar o cotidiano no interior dos domicílios coloniais é uma tarefa difícil, porque são extremamente raros registros sistemáticos. Para Santos Filho (2012, p. 20), segundo o qual “como que armado de uma grossa lente de aumento, procurei focalizar, em ponto grande, um pequeno centro nordestino do Brasil antigo”, o *Livro do Gado* e o *Livro de Razão* “são a porção mais valiosa do opulento arquivo do sobrado do Brejo” (LACERDA, SANTOS, 2022, p. 14).

Os capítulos e artigos somam, atualmente, 286 publicações. É uma vasta produção, divulgada também em eventos locais, regionais, nacionais e

4 A edição fac-similar do *Livro do Gado* e do *Livro de Razão* foi feita segundo o método de fotografia Lapelinc (SANTOS, BRITO, 2014).

5 O NELP publicará, em breve, a Série *Documentos para a História Linguística do Brasil Colônia*, organizada por Carneiro, Lacerda e Lose, que reúne, sobretudo, documentos editados como parte de dissertações de mestrado e teses de doutorado (já concluídas ou em andamento), orientadas por Carneiro, Lacerda e Lose; neste momento, encontram-se em editoração, na UEFS Editora, os três primeiros volumes da Série.

internacionais, entre eles 16 eventos organizados pelo NELP, como O I Ciclo de Palestras do NELP: diálogos interinstitucionais e O I Encontro do CE-DOHS: língua, história e tecnologia, ambos em 2020, na modalidade remota, durante a pandemia de COVID 19. A divulgação das produções e eventos do núcleo é feita também em suas redes sociais, como dito na seção 1.4.

1.3 Rede de pesquisa

O NELP participa, desde sua criação, de uma rede de pesquisa forte, dialogando com pesquisadores de diferentes instituições brasileiras, como, na Bahia, a Universidade do Estado da Bahia (UNEB), Universidade do Sudoeste da Bahia (UESB) e a UFBA, e, fora da Bahia, a Universidade Federal de Pernambuco (UFPE), Universidade Federal de Sergipe (UFS), Universidade Estadual de Campinas (UNICAMP), Universidade do Estado de São Paulo (USP), Universidade Federal do Rio de Janeiro (UFRJ), Universidade Federal de Minas Gerais (UFMG) e Universidade de Brasília (UnB). Fora do Brasil, estabelece parceria com a Universidade de Lisboa (UL), por meio de atuação no projeto Pombalia: Pombal Global, coordenado pelos professores José Eduardo Franco e Pedro Calafate.

Com a UNICAMP foi firmado convênio, com Termo Aditivo de Transferência de Tecnologia, desde 2012; uma parceria formalizada entre o projeto Corpus Histórico do Português TychoBrahe (UNICAMP)⁶, coordenado pela professora Charlotte Marie Chambelland Galves, e o projeto Corpus Eletrônico de Documentos Históricos do Sertão (CE-DOHS/NELP/UEFS), coordenado pelas professoras Zenaide de Oliveira Novais Carneiro e Mariana Fagundes de Oliveira Lacerda, apresentado na seção 2. Diferentes planos de trabalho têm sido desenvolvidos, por meio desse convênio; planos de edição em linguagem XML, com modernização, e planos de anotação sintática.

6 Disponível em: <<https://www.tycho.iel.unicamp.br/corpus/>>.

1.4 Conexão NELP

O NELP se comunica com a comunidade interna da UEFS e com a comunidade externa, também por meio de suas redes sociais, desde 2017, quando a Coordenação do núcleo passou a investir, de forma mais ativa, na divulgação de suas produções e eventos.



Figura 6: Página inicial do site institucional do NELP⁷.

Fonte: <<https://nelp.uefs.br/>>

Além do site institucional, no endereço <https://nelp.uefs.br/>, cuja página inicial está apresentada acima, o NELP está no Instagram, no Facebook, no Twitter e no YouTube⁸.

7 O site do NELP está, neste momento, passando por atualizações, como parte das comemorações pelos seus 25 anos, com um menu especial sobre seu aniversário, serviço realizado com recursos do edital interno UEFS 001/2021/Termo de Outorga nº 048/2021.

8 Instagram: @nucleonelp;
Instagram: @ce_dohs;
Facebook: Ce-dohs;
Twitter: @ce_dohs;
YouTube NELP: @nelpnucleodeestudosemlingu8083
<<https://www.youtube.com/@nelpnucleodeestudosemlingu8083/playlists>>
YouTube CE-DOHS: @ce-dohs5910 <<https://www.youtube.com/channel/UCeh6wC1GzIzslcBHCBwKLxQ?app=desktop>>.

No Instagram, por exemplo, o NELP e o CE-DOHS, juntos, têm 1.576 seguidores, um número significativo, tratando-se de páginas acadêmicas, com conteúdo sobre Linguística Histórica, História do PB e Ensino de Língua Portuguesa.

Esse diálogo com a comunidade interna da UEFS e externa, em diferentes plataformas e linguagens, é muito importante, sobretudo atualmente, quando estamos todos conectados com o mundo, também por meio da rede mundial de computadores.

2 O banco de dados sociolinguísticos do NELP

Como afirmou Mattos e Silva (2004, p. 61), a busca pela reconstrução da história do PB “será trabalho para muitas mãos, durante muito tempo”. E esse trabalho inclui, fundamentalmente, a constituição de *corpus*, onde buscamos os dados, que iluminam a linguística teórica. Conforme Bacelar do Nascimento (2004, p. 1), “O uso de *corpora* permite a realização de descrições linguísticas de base empírica e promove, com isso, a discussão de questões teóricas solidamente fundamentadas.”

O NELP tem, desde sempre, o compromisso com a formação de banco de dados, com o necessário controle sócio-histórico – a qual está entre as agendas principais do PHPB –, reconhecendo que

A formação de banco de dados, como o Corpus Eletrônico de Documentos Históricos do Sertão (CE-DOHS), do Núcleo de Estudos de Língua Portuguesa (NELP) da Universidade Estadual de Feira de Santana (UEFS), exige grande esforço e tempo dedicados à sua metodologia de organização: prospecção documental em fontes confiáveis; caracterização sócio-histórica de acervos; edições em diferentes formatos; processamento, armazenamento e disponibilização de dados. Como se vê, um trabalho de pesquisa que envolve diferentes etapas e uma equipe com formação especializada em campos diversos, como a Filologia, a História Social da Cultura Escrita, a Linguística Computacional, entre outros.” (SANTIAGO *et al*, 2021, p. 314)

O banco Documentos Históricos do Sertão (DOHS) – que ganhou, a partir de 2010, uma versão eletrônica, o CE-DOHS (conferir seção 3.1) – tornou-se um banco representativo da polarização e pluralidade do PB (LUCCHESI, 1994)⁹, cobrindo, dentro do possível, fases da periodização da história sociolinguística da língua (LUCCHESI, 2017), com textos, em sua maioria, inéditos. O núcleo sempre procurou reunir, na constituição de *corpora* diacrônicos, documentação variada, sendo possível que o pesquisador personalize seu *corpus*, segundo seu interesse.

O CE-DOHS permite que se personalize o *corpus*, de acordo com o interesse do pesquisador, que, entre inúmeras possibilidades, pode optar em o separar, no que concerne ao autor, por: etnia (indígenas, brancos, negros do Brasil, mulatos, mamelucos e pardos, entre outras); nível de escolarização e habilidade/inabilidade com a escrita; sexo; profissão; estratificação social; data e local de nascimento do autor (naturalidade e nacionalidade); e, no que concerne ao documento, por: data e local de escrita, meio urbano e meio rural, para quem e a quem foi destinado (CARNEIRO; LACERDA, 2023, Início).

O NELP não tem, entretanto, o *corpus* ideal nem trabalha com esse conceito; o que oferecemos é um banco, com lacunaridade¹⁰, para uma aproximação da história linguística dos diferentes grupos envolvidos nessa complexa formação do Brasil, sabendo que, no que diz respeito aos grupos sociais subalternos, as lacunas são ainda maiores, haja vista, como observa Lobo (2001, p. 109), que “estes raramente deixaram testemunhos autógrafos”, a escrita de sua história linguística pautando-se, em muitos aspectos, “em

9 Lucchesi (2017, p. 376) aponta um nivelamento linguístico na 4ª fase (1930 até hoje) de sua proposta de periodização sociolinguística do Brasil, assim caracterizada: “Industrialização e urbanização promovem a difusão da norma urbana culta, apagando as marcas do contato na norma popular; a norma culta se afasta do padrão lusitanizado; porém, concentração de renda mantém a polarização sociolinguística e fomenta o preconceito linguístico.”

10 Conferir Paixão de Souza (2006) sobre a tradição documental dos estudos históricos sobre as línguas e a questão da lacunaridade.

uma reconstrução a partir de “indícios”, e, necessariamente, do presente em direção ao passado”.^{11 12} Como disse Mattos e Silva (2008b), trata-se de um trabalho de arqueologia da linguagem.

Um banco de dados sociolinguísticos é o que o NELP disponibiliza a pesquisadores interessados, permitindo análises histórico-diacrônicas do português brasileiro. Com objetivos bem definidos, de cobertura da história sociolinguística do Brasil, os pesquisadores do NELP constituíram, a partir de um intenso trabalho de prospecção documental, atrás de textos remanescentes, o seu *corpus* de trabalho, pois, conforme ressalta Mattos e Silva (1989, p. 15),

Em qualquer estudo que se volte para uma análise de manifestações linguísticas em uso, o *corpus* sob análise, em geral, é constituído pelo linguista de acordo com os objetivos de sua pesquisa. O *corpus* será então, necessariamente, um corte intencional sobre a totalidade de uma língua; a sua data, o seu lugar, os seus informantes serão identificados e definidos. Quando tratamos de um estado de língua já passado, o corpus sobre que se desenvolverá a análise, por sua própria natureza, já está constituído. O filólogo, o linguista, o <<poéticien>> determinarão os limites de seu campo de observação sobre documentação preexistente. (MATTOS E SILVA, 1989, p. 15)

11 Para Lobo (2009, p. 109), um *corpus* diacrônico ideal para estudo da história social e linguística do Brasil seria: “(i) Subcorpus 1: as variedades do português europeu transplantadas para o Brasil; (ii) Subcorpus 2: as variedades do português falado como segunda língua pelos aloglotas; (iii) Subcorpus 3: as variedades do português brasileiro que paulatinamente se iam constituindo: 3.1) as variedades cultas – supostamente mais unitárias e descendentes diretas das variedades do português europeu – e 3.2) as variedades populares – supostamente mais diversificadas e descendentes diretas das variedades do português como segunda língua.”

12 No CE-DOHS, as coleções documentais representativas das normas socialmente estigmatizadas do PB são: Correspondências Amigas; Cartas em Sisal; Cartas Marienses, do século XX. Sobre as Cartas em Sisal, editadas por Santiago (2012), foi publicada uma coletânea de estudos morfossintáticos, organizada por Santiago, Lacerda e Carneiro (2023), com financiamento do Programa de Auxílio Interno aos Programas de Pós-Graduação *Stricto Sensu* (AUXPPG/UEFS).

Além dos documentos – a maioria cartas¹³ –, estão disponíveis no site (<http://www.uefs.br/cedohs/view/sobre.html>), com acesso livre e gratuito, em consonância com os princípios da Ciência Aberta, seus metadados e os metadados dos escreventes/informantes, também informações sobre o processamento dos documentos (SANTIAGO *et al*, 2021)¹⁴. Tudo feito de forma sistemática, etapa a etapa, desde a prospecção documental até a disponibilização dos materiais, dando o núcleo contribuições importantes à discussão sobre o tratamento metodológico à constituição de *corpora*. Nos procedimentos metodológicos, o NELP sempre considera o *corpus* no seu tempo (e, nos estudos sociolinguísticos, o *corpus* nos diz o que tem a oferecer).

Na descrição do perfil sociocultural dos escreventes, para a reconstituição da sócio-história do PB, entendendo a escrita a partir de quem a escreveu em um determinado contexto sócio-histórico, o NELP vai além das categorias tradicionais de perfil, investindo esforços para o que Lopes *et al* (2010, p. 242) denominaram de “cruzamento de perspectiva”:

Acreditamos no potencial analítico de um cruzamento de perspectivas, que relacione a trajetória de vida dos sujeitos (que vai mais além das categorias tradicionais de perfil), o contexto de produção dos textos (em que momento foi escrito, o que foi escrito, para quem foi escrito, em que condições e com que finalidades foi escrito) e o mapeamento e descrição

13 Conforme Lacerda, Carneiro e Santiago (2016, p. 131), “A maior parte dos documentos do DOHS, datados e localizados – que hoje se encontram também em versão digital no CE-DOHS – são cartas manuscritas, dos séculos XIX e XX (1084 cartas, 422 remetentes), editadas sobretudo por Carneiro (2005), que investiu na busca e na organização de acervos documentais que pudessem contribuir para o processo de reconstrução sócio-histórica do PB, em um trabalho de investigação grandioso, percorrendo diversos arquivos, e publicadas em 2011, pela Editora UEFS, na obra, com três volumes, organizada por Carneiro, Lacerda, Almeida, Santiago e Oliveira, intitulada *Cartas brasileiras: coletânea de fontes para o estudo do português*.” Em 2022, foi publicado o volume 4 da Coletânea (BRITO; LACERDA, 2022) e, em 2023, o volume 5 (CARNEIRO, TUY BATISTA; ALMEIDA, 2023).

14 O *corpus* encontra-se também compartilhado na página *online* do PHPB Bahia: <<https://sites.google.com/site/corporaphbba/?pli=1>>.

das redes de escrita (diálogos estabelecidos e possíveis interlocutores). Isto seguramente nos permitirá localizar a produção escrita de um indivíduo num contexto de produção mais amplo, o que, por sua vez, nos garante a possibilidade de uma conceituação alterna de perfil sociocultural. (LOPES *et al*, 2010, p. 242)

É o caminho que o NELP vem percorrendo, tentando, o máximo possível, abordagens sobre o indivíduo, as quais considerem esse contexto de produção mais amplo. Fazendo o controle sócio-histórico da documentação, o CE-DOHS descreve sobre o “quem” e “para quê” (PETRUCCI, 2003), além de descrever, no âmbito da paleografia tradicional, o “quando” e “onde”. São variáveis que, segundo Mattos e Silva (2004), amenizam, com o recurso à historiografia, o problema dos “maus dados” (LABOV, 1982) com que trabalha o pesquisador da Linguística Histórica.

Uma grande equipe de especialistas, dedicados, “Como quixotes ou como loucos, ou apenas como brasileiros interessados em compreender um aspecto fundamental da sua história pregressa” (MATTOS E SILVA, 2004, p. 67), ao desafio de reconstrução, por aproximação, de uma história do PB. E, se tudo mais pode passar, se as teorias podem ficar para trás, os dados não passam, as descrições ficam, como sempre nos lembrava, nas suas aulas no Instituto de Letras da UFBA, a pró Rosa... Então, aqueles que constituem *corpus*, com grande esforço, fazem um trabalho necessário e valoroso:

O trabalho com *corpus* não é “commodo” – pelo contrário, é por vezes penoso, sempre curioso, às vezes divertido (...) Sem dúvida é um trabalho necessário, como base para a reconstrução do passado lingüístico do português que aqui se formou, o português brasileiro (MATTOS E SILVA, 2004, p. 120).

E, todo o tempo, como não poderia deixar de ser, com a preocupação com a autenticidade, porquanto, como ressalta Maia (2012, p. 538),

Sob o ponto de vista qualitativo, a autenticidade é uma exigência decisiva: se se pretende utilizar os materiais para o estudo de história da língua, as edições devem satisfazer as necessidades de uma linguística histórica

empírica e, por esse motivo, devem reflectir fielmente as características linguísticas dos manuscritos, uma vez que nelas se reflectem as marcas da variação da língua da época. (MAIA, 2012, p. 538)

A importante parceria da Linguística Histórica com a Filologia Textual e a Paleografia de Análise, à qual já nos referimos outras vezes aqui, esse cruzamento de olhares como condição *sine qua non* para os estudos histórico-linguísticos.

Os critérios de edição semidiplomática adotados pelo NELP são, normalmente, os do PHPB (CASTILHO, 2020), mas também, sobretudo no tratamento dos documentos do período colonial, adotamos os critérios de transcrição do Centro de Estudos, Pesquisa e Documentação Paleográfica (CEPEDOP/<<https://www.memoriaarte.com.br/>>).

A depender das especificidades dos documentos, esses critérios sofrem adaptações, sempre esclarecidas ao leitor (conferir, por exemplo, dissertação de Santos (2019) e tese de Silva (2021)). São edições semidiplomáticas, tendo em vista que são feitas algumas intervenções: sempre mantidas as características originais dos textos, mas desenvolvidas as abreviaturas e marcadas, com algum elemento, as letras antes ausentes; também empregados caracteres atualizados (LOSE, 2017, p. 74).

No processo de edição – o *corpus* nas mãos da Filologia –, elaboramos, inicialmente, o quadro escriptográfico do escrevente, na descrição intrínseca do documento. Os aspectos intrínsecos do documento, de acordo com Lose *et al.*(2009),

são definidos aqui como aquelas características ainda não ligadas à “língua”, mas sim às peculiaridades “ortográficas” de cada scriptor. É importante fazer uma ressalva para o fato de que “ortografia”, neste contexto, não deve ser pensada como a escrita correta, mas sim como a forma de escrever e de dispor e combinar os grafemas, criando, desta forma, fatos linguísticos a serem analisados. (LOSE *et al.*, 2009, p. 47).

O estudo descritivo dos grafemas e das peculiaridades da escrita de cada escrevente são indispensáveis para a análise de documentos antigos,

porquanto, a partir desse estudo minucioso, é possível conhecer, com mais detalhes, a mão que escreve; o trabalho paleográfico é imprescindível para fazer uma boa decodificação da escrita de um texto antigo (CAMBRAIA, 2005). Além das peculiaridades gráficas, observamos, também, na descrição intrínseca, o conteúdo dos documentos. Essa observação e estudo das peculiaridades gráficas são fundamentais para o processo de edição do texto.

Sobre os *corpora* orais do banco, são, em grande parte, produto do projeto A Língua Portuguesa no Semiárido Baiano, já aqui referido. Do projeto Em Busca das Raízes do Português Brasileiro, criado pela professora Eliana Pitombo Teixeira, no âmbito do NELP e, atualmente, coordenado pela professora Silvana Silva de Farias Araujo, são as amostras de fala de Luanda-Angola (UEFS), que estão em processo de revisão, de maneira que possam também, em breve, serem disponibilizadas no banco. Há outras amostras, já disponíveis no banco, obtidas por meio de parceria com o Programa de Estudos sobre o Português Popular de Salvador (PEPP-UNEB), coordenado pela professora Norma da Silva Lopes, e o Projeto Norma Linguística Urbana Culta do Brasil (NURC/Salvador/UFBA), coordenado pela professora Jacyra Andrade Mota. Também por meio de parceria, serão, em breve, disponibilizadas no banco as amostras do Projeto Estudos Linguísticos e Históricos do Sertão (ELIHS-UNEB), coordenado pela professora Dayane Moreira Lemos, que defendeu, recentemente, no PPGEL da UEFS, a tese, orientada pela professora Mariana Fagundes de Oliveira Lacerda, intitulada *ELIHS, um projeto para além do litoral: da constituição de corpus oral de comunidades afro-brasileiras à análise estatística da variação na concordância nominal de número na variedade da comunidade afro-brasileira de Volta do Angico (BA)* (LEMOS, 2023).

Cada projeto adota suas normas de transcrição próprias; os projetos mais recentes, como o de Lemos (2023), baseando-se também nas normas dos projetos mais antigos; a autora adora a chave de transcrição ortográfica dos inquéritos realizados no âmbito do Projeto Vertentes do Português Popular do Estado da Bahia (<http://www.vertentes.ufba.br>).

No site do CE-DOHS, estão disponíveis a edição semidiplomática dos manuscritos e impressos¹⁵ e a edição modernizada, em linguagem XML; também as amostras de fala, sua transcrição e versão modernizada, em XML.

Segundo Shepherd *et al.* (2012, p. 11),

A ideia de coligir coleções de textos naturais com o objetivo de os submeter à análise linguística remonta ao trabalho dos estruturalistas norte-americanos da década de 1950, tais como Harris (1951) e Fries (1952). Com o Brown Corpus (Francis e Kucera, 1954), surgiria o primeiro corpus eletrônico compilado para este fim. Embora até hoje este *corpus* seja largamente utilizado, na altura praticamente não existiam textos escritos em formato digital, os computadores eram máquinas enormes e caras, que ocupavam salas inteiras, e os programas informáticos demoravam horas e até dias a correr. (SHEPHERD *et al.*, 2012, p. 11)

O CE-DOHS – versão eletrônica do DOHS¹⁶ – soma-se aos novos *corpora* eletrônicos, como o Corpus Tycho Brahe, da UNICAMP, na era das Humanidades Digitais, rótulo que, segundo Castro (2020, p. 2), inclui “tudo que signifique a aplicação de computação e de tecnologias digitais ao universo das humanidades. O pano de fundo é a expansão de big data, conjuntos de dados grandes demais para serem analisados por formas tradicionais de pesquisa.”¹⁷ E, nesse cenário, dá-se o “crescimento vertiginoso da Linguística de Corpus” (SARDINHA, 2000, p. 323), que, de acordo com Sardinha (2007, p. 325), “ocupa-se da coleta e da exploração de *corpora*, ou do conjunto de dados linguísticos textuais coletados criteriosamente, com o propósito de servirem para a pesquisa de uma língua ou variedade linguística”.

15 Conferir a obra, organizada por Carneiro e 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)*, que traz documentos impressos, do século XX, os mesmos do período que estão disponíveis no CE-DOHS. No site, há documentos impressos do século XIX também, estes publicados na obra, organizada por Barbosa e Lopes (2006), intitulada *Críticas, queixumes e bajulações na Imprensa Brasileira do século XIX: cartas de leitores*.

16 Neste momento, o site do projeto CE-DOHS está passando por atualizações, com a inserção de novos documentos, tanto da fase 1 como, principalmente, da fase 2, com recursos do Edital UEFS 001/2021 – Auxílio Financeiro a Projetos de Pesquisa e Inovação.

17 Conferir Schreiban *et al.* (2004).



Figura 7: Página inicial do site CE-DOHS.

Fonte: <http://www.uefs.br/cedohs/>

Com mais de 10 anos ativo na rede mundial de computadores, o CE-DOHS reúne uma extensa base documental, resultante, como já dissemos aqui, de árduo trabalho, organizada em dois conjuntos: conjunto 1 – composto por textos escritos entre 1808 e 2000, por indivíduos nascidos no Brasil, a partir de 1724, e por amostras de fala de brasileiros, gravadas na década de 90 do século XX, na Bahia; conjunto 2 – composto por manuscritos produzidos entre 1640 e 1808 por diferentes populações nascidas no Brasil, a partir de 1590, e em processo de edição no mesmo formato do conjunto 1. O conjunto 2 encontra-se em elaboração, tratando-se de um *corpus* mínimo, dada a raridade de documentos. Adicionalmente, há um conjunto de textos escritos no Brasil por portugueses, nos primeiros 150 anos da colonização (CARNEIRO; LACERDA, 2023, Apresentação).^{18 19}

18 Sobre, especialmente, a fase 1 do CE-DOHS, conferir Carneiro e Lacerda (2019).

19 Sobre a fase 2, a representatividade e potencialidades do *corpus* colonial do CE-DOHS, conferir Cardoso, Carneiro e Lacerda (2021).

Aliando a Antiga e a Nova Filologia²⁰, o CE-DOHS apresenta os documentos em diferentes versões de edição; além dos fac-símiles – que dão total transparência ao labor filológico realizado –, a edição semidiplomática e a edição modernizada, em XML.

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).

A edição modernizada, com uso do eDictor (PAIXÃO DE SOUSA; KEPLER; FARIA, 2009), permite a anotação sintática, que se encontra em andamento no âmbito do CE-DOHS, por meio de convênio, especialmente com a UNICAMP, interessada no uso do banco para estudos sintáticos de natureza gerativista. Todas as intervenções feitas no texto, com uso do eDictor, ficam visíveis ao leitor, o que possibilita seu controle e mapeamento, garantindo a recuperabilidade das formas originais. O referido programa dispõe, atualmente, para o processo de modernização dos textos – concernente apenas a intervenções de natureza gráfica, por conseguinte uma modernização com baixo grau de intervenção –, de onze tipos de etiquetas: junção, segmentação, sobrescrito, ilegível, rasurado, subscrito, tachado, pontuação, expansão, padronização e modernizado.

No site do CE-DOHS, no *menu* Coleções Documentais/Edições em Diferentes Formatos, é possível acessar as diferentes versões de edição, das três categorias maiores de documentos: Manuscritos, Impressos, Amostras de Fala. Somam-se 20 as coleções documentais da fase 1 (1823-2000), já concluída, e 24 as coleções da fase 2, iniciada em 2017 e em andamento.

20 Conferir Crane *et al.* (2008).

Os interessados podem visitar o site do CE-DOHS e navegar, fazendo as buscas que desejar. É possível usar, para fazer buscas de dados, a ferramenta *E-Corp*, desenvolvida para aplicação inicial no CE-DOHS:

Com o crescente desenvolvimento da construção de banco de dados eletrônicos, surge a necessidade de criação de ferramentas que auxiliem na exploração de documentos em formato XML. Dessa maneira, na tentativa de otimizar o contato inicial do pesquisador com os *corpora*, foi desenvolvida a ferramenta *E-Corp*, que torna a busca nos bancos de dados mais rápida e confiável, além de permitir a exploração dos acervos, ajudando na construção de *subcorpora*, já que é possível filtrar as informações sobre o documento a partir dos metadados de cada documento (SOUZA *et al*, 2018, p. 14).

Utilizando essa ferramenta, o consultante pode montar seu *corpus* de estudo, conforme seus interesses.

Em breve, será possível também acessar, no site, cuja documentação (aproximadamente cinco mil documentos) reúne 2,3 milhões de palavras – já tendo ultrapassado as fronteiras dos sertões –, os locais de produção e/ou recolha, relativos a coleções documentais, por meio de fotografias ilustrativas em modelagem em 3D, resultantes de recriação de fotografias reais ou de descrição dos referidos locais.²¹

A edição modernizada, em linguagem XML, com uso do eDictor, impôs desafios à equipe do NELP, que aprendeu a manusear a ferramenta computacional, acessando os tutoriais *online* (<https://humanidadesdigitais.org/edictor/>) e participando de oficinas e minicursos, a maioria organizados pelo núcleo, nos eventos promovidos, sobre Humanidades Digitais e Linguística de Corpus.

Como o CE-DOHS apresenta documentos de gêneros diversos, há, às vezes, desafios diante dos quais alguns deles, tendo em vista suas especificidades, colocam o eDictor. É o caso dos livros manuscritos da

21 Conferir Tuy Batista *et al* (2023). Trabalho de pesquisa financiado com recursos do Edital UEFS 001/2021 – Auxílio Financeiro a Projetos de Pesquisa e Inovação.

Família Pinheiro Canguçu, já aqui referidos, o *Livro do Gado* e o *Livro de Razão*, conforme demonstram Lacerda *et al* (2023), em texto intitulado *Edição modernizada dos dois livros manuscritos do Brejo do Campo Seco, BA (XVIII-XIX): desafios da nova filologia*.

É um longo percurso até aqui, da prospecção documental à disponibilização de dados. A transição do DOHS ao CE-DOHS ainda está em curso; não podemos ainda afirmar que se trata o CE-DOHS de um banco eletrônico prototípico, mas está no caminho para tornar-se um, na medida em que pesquisadores parceiros, interessados no banco, investirem esforços, por exemplo, no trabalho de anotação sintática dos textos – que permite a busca de construções anotadas, a partir de perguntas de investigação (GALVES, 2019) –, o que já está sendo feito, especialmente por Charlotte Galves e Williane Corôa²². O CE-DOHS é hoje um banco

22 A anotação sintática – que é feita segundo a teoria Gerativa, principalmente – não se trata de uma etapa de trabalho do CE-DOHS, mas de um dos usos possíveis que o pesquisador pode fazer do banco, como Charlotte Galves, gerativista, professora da UNICAMP, e Williane Corôa, gerativista, professora da Universidade do Estado da Bahia (UNEB), estão fazendo com seus orientandos, como parte de planos de IC, de Mestrado e de Doutorado (conferir artigo intitulado *Relações entre a constituição de corpora diacrônicos do português brasileiro e os princípios da Ciência Aberta*, de Cardoso, Lacerda e Carneiro (no prelo). O CE-DOHS pode ser explorado de acordo com a concepção de língua do pesquisador e, dela decorrentes, com suas opções teórico-metodológicas, ficando o banco disponível a interessados nos diferentes caminhos da Linguística Histórica: intrassistêmico (objeto a-histórico) e extrassistêmico (objeto histórico), configurando-se o CE-DOHS, ressaltando, em um *corpus* especial para os que consideram a historicidade das línguas, as relações entre linguagem e sociedade, para os que trabalham com um objeto histórico, haja vista a riqueza de metadados, com cruzamento de perspectivas, apresentada no site. O CE-DOHS/NELP/UEFS defende a pluralidade teórica, que, como disse Faraco (2005, p. 92), “não é apenas um fenômeno desejável, mas uma necessidade lógica. Como não temos o dom da onisciência, nem o poder de apreensão global instantânea do mundo, nossas aproximações científicas do real são sempre parciais: fazemos recortes nele, construindo nossos objetos de estudo, e formulamos hipóteses explicativas para esses recortes.” Então, as etapas de prospecção documental, controle sócio-histórico, descrição de metadados, edição semidiplomática, edição modernizada e disponibilização de dados são da própria constituição do banco em si, no âmbito da UEFS; a etapa de anotação sintática, todavia,

digital²³, que oferece diversos documentos em diferentes versões; da fase 1, já se encontram as coleções documentais, a maioria, totalmente modernizadas, em linguagem XML; da fase 2, iniciada há pouco tempo, não há, no banco, ainda, nenhum texto modernizado, porque, no momento, a equipe do projeto está dedicada ao trabalho de edição semidiplomática, ao estudo sócio-histórico e estudo linguístico dos materiais levantados. Um longo percurso, sempre com metodologias criteriosas e claras, em busca de respostas a perguntas como as seguintes, entre outras, sintetizadas na apresentação da equipe PHPB-Bahia (Sertões), no X seminário do PHPB, na UFS, em 2019:

- a) Quais são as dinâmicas de mudança do português no Brasil?
- b) Quais são os aspectos da língua afetada pela mudança?
- c) Quais são as parametrizações envolvidas?
- d) Como constituir um *corpus* sistemático que permita capturar a história complexa de formação do português brasileiro?
- e) Como superar as dificuldades do recuo ao período anterior ao século XVIII?
- f) Como capturar a natureza peculiar do PB, uma língua marcada pelo contato?
- g) É possível rastrear os efeitos do contato em textos escritos?

vem sendo executada na medida do interesse de pesquisadores parceiros na exploração dos dados na perspectiva gerativista. São buscas possíveis que se podem fazer no *corpus*, com uso de diferentes ferramentas, sejam ferramentas para busca de dados sintaticamente anotados, sejam ferramentas para buscas lexicais. Dessa forma, o investimento de esforços do CE-DOHS/NELP/UEFS é para a constituição de *corpus*, com o devido controle sócio-histórico e rigor filológico e paleográfico, contribuindo com os que se interessam pelo estudo da história social linguística do português brasileiro, em diferentes níveis de análise e dentro de diferentes quadros teóricos. O banco passa a ser prototipicamente eletrônico, na medida em que pesquisadores interessados vão preparando os materiais para as buscas que desejam fazer, usando diferentes ferramentas, de acordo com seus objetivos/perguntas.

- 23 Sobre o conceito de texto na esfera digital, conferir Paixão de Souza (2013) e Marquilhas e Hendrickx (2016), entre outros.

Na sua vasta produção bibliográfica, o NELP traz resultados de pesquisas que, sem dúvida, colaboram para elucidar questões como as supracitadas. E o próprio banco de textos do núcleo, do modo como ele se constitui e se apresenta – com base também no alinhamento teórico e metodológico proposto pela Sociolinguística Histórica (ROMAINE, 1982; HERNÁNDEZ-CAMPOY; CONDE-SILVESTRE, 2012) –, responde a problemas enfrentados para a reconstrução da história social linguística do PB.²⁴

O estudo sócio-histórico é realizado antes e após a prospecção documental, sendo etapas essenciais no processo de constituição de *corpus*. Na etapa de prospecção de material, é feito o estudo da realidade sócio-histórica, indo atrás de materiais que nos ajudem a preencher as lacunas existentes; e, uma vez localizados esses materiais, o controle sócio-histórico é feito, garantindo a legitimidade do *corpus*. Como resultado desse estudo, o NELP possui diversos trabalhos publicados, especialmente considerando a região do Semiárido baiano; podemos citar, entre outros: Carneiro e Almeida (2006; 2008; 2011); Lacerda, Araújo e Carneiro (2018); Oliveira e Araújo (2018); Carneiro, Lacerda e Almeida (2020); Carneiro *et al* (2020).

Os estudos linguísticos, até o presente, realizados pelo NELP, explorando o seu banco de dados, são muitos, como, por exemplo, os estudos de natureza variacionista, explorando *corpus* oral, reunidos por Almeida *et al* (2016) em coletânea intitulada *Variação Linguística em Feira de Santana-Bahia*, publicada pela UEFS Editora. Também o estudo do sistema de tratamento, analisando diferentes coleções documentais epistolares, em parceria interinstitucional, coordenado pela professora Célia Regina dos Santos Lopes, da UFRJ (MARTINS *et al*, 2015; ANDRADE; CARNEIRO; LACERDA, 2016; LACERDA; CARNEIRO; OLIVEIRA, 2016; TUY BATISTA; CARNEIRO; LACERDA, 2017; LOPES *et al*, 2018; FREITAS;

24 Conferir capítulo de Santiago e Brito (2023, p. 123-135), intitulado *Pesquisas em Sociolinguística Histórica na Bahia: alguns percursos e possibilidades*, publicado no livro *A Sociolinguística no Nordeste: diferentes abordagens*.

LACERDA; CARNEIRO, 2018; SANTOS; LACERDA; CARNEIRO, 2019a; SANTOS, LACERDA; CARNEIRO, 2019b; SANTOS; LACERDA; CARNEIRO, 2022). Ainda, entre outros, o estudo, de natureza morfossintática, da coleção documental *Cartas em Sisal*; trata-se de 10 trabalhos reunidos por Santiago, Lacerda e Carneiro (2023) em coletânea publicada pela Editora Pontes, intitulada *Cartas em Sisal: estudos morfossintáticos*.

No âmbito do NELP, a documentação é explorada, mesmo quando ainda não editada filologicamente, na íntegra; o pesquisador edita apenas pedaços do texto que interessam em seu estudo, constituindo um exemplário, como fez Cardoso (2020) em sua dissertação de Mestrado, orientada pela professora Zenaide de Oliveira Novais Carneiro e defendida no PPGEL/UEFS.

O NELP, ao longo desses 25 anos, contou muitas histórias do PB, por meio da rica documentação que pôde reunir, “com amor, paixão, rigor e paciência” (MATTOS E SILVA, 2013, p. 678), como deve ser o trabalho de constituição de *corpus*, contribuindo para a curadoria do patrimônio histórico. Contou e conta muitas histórias, fazendo sua história na UEFS e no Semiárido baiano.²⁵

3. NELP na sala de aula

Com um banco de textos que se tornou referência no Brasil e fora do Brasil para estudo da história do PB e diversos estudos sócio-históricos e linguísticos dessa documentação, o NELP não poderia deixar de assumir também o compromisso com a Educação, fazendo chegar, de alguma forma, à sala de aula essa riqueza de dados e valiosa produção, financiada, em grande

25 Marquilhas e Hendrickx (2016) discorrem sobre *Avanços nas Humanidades Digitais* e citam, entre os projetos de *corpora* históricos reconhecidos internacionalmente, o CE-DOHS (p. 21). O capítulo faz parte do livro, *Manual de Linguística Portuguesa*, organizado pelas professoras Ana Maria Martins e Ernestina Carrilho, publicado pela De Gruyter.

parte, com recursos públicos, por meio da submissão de propostas a editais de financiamento de pesquisa.

Desde a sua criação, em 1997, pelas professoras Zenaide de Oliveira Novais Carneiro e Norma Lucia Fernandes de Almeida, já estava, entre as agendas de trabalho do NELP, a agenda educacional. Mas, dada a necessária dedicação, ao longo de muitos anos, à constituição do banco de textos e aos estudos sócio-históricos e linguísticos, a agenda educacional não pôde receber, no âmbito do núcleo, a atenção desejada e necessária. Somente, portanto, a partir de 2017, quando foi criado pelas professoras Mariana Fagundes de Oliveira Lacerda e Zenaide de Oliveira Novais Carneiro o projeto NELP na Sala de Aula: diálogos entre pesquisa, ensino e extensão, o núcleo passou a investir, sistematicamente, esforços para desenvolver essa agenda. Em 2017, a fase 1 do CE-DOHS já estava bastante adiantada, com várias coleções documentais disponíveis na página *online*; também em 2017, as primeiras dissertações elaboradas no âmbito do Mestrado Profissional em Letras (PROFLETRAS) da UEFS, e orientadas por essas duas professoras, já haviam sido defendidas. Então, o momento foi oportuno para dedicar mais atenção à exploração do CE-DOHS na interface educacional. E esse objetivo ficou sendo um dos principais da coordenação do NELP, assumida pela professora Mariana Fagundes de Oliveira Lacerda, em 2017, empenhando-se o núcleo, a partir daí, cada vez mais, para um diálogo com a Educação Básica, fazendo chegar às escolas – em formato de produtos educacionais, como cadernos de atividades, audiovisuais, blogs etc. – resultados de pesquisas acadêmicas, como parte do projeto NELP na Sala de Aula e também em parceria com a Ação Linguística na Escola, do PPGEL da UEFS.²⁶

26 Conferir mesas redondas da Ação Linguística na Escola, no canal do PPGEL, no YouTube. Disponível em: <https://www.youtube.com/channel/UCBF6HpKXIZY4QA0mv9dA6ZA>. Acesso em: 22 set. 2022.



Figura 8: Cartaz de divulgação do projeto NLP na Sala de Aula.

Fonte: <https://nlp.uefs.br/nlp-na-sala-de-aula/>.

Não é consensual que a Linguística tenha um compromisso necessário com a Educação. Castilho (2003), por exemplo, considera fundamental o diálogo entre a Linguística e a Educação; Borges Neto (2003), por outro lado, afirma que a Linguística, em sentido estrito, não tem um compromisso com a Educação, embora tenha permitido avanços muito grandes na área educacional. Mas não se nega, de nenhum parte, que os linguistas podem contribuir tanto para o ensino de língua materna quanto para o ensino de línguas estrangeiras. E é o que o NLP tem também procurado fazer, dar contribuições ao ensino de português como língua materna, questão sobre a qual se debruçou também Mattos e Silva (MATTOS E SILVA, 1996; MATTOS E SILVA 2004b).

Até o presente, o núcleo já realizou diversos eventos, vinculados ao projeto NLP na Sala de Aula. A última edição, intitulada NLP na Sala

de Aula: linguística, prosa e música, aconteceu, na modalidade remota, em 2022, reunindo estudantes dos cursos de Letras da UEFS e professores da Educação Básica.

Também o NELP vem publicando capítulos e artigos com reflexões voltadas ao ensino de português. Recentemente, em 2023, foi publicado o livro *Ensino de português: fonologia, morfologia e sintaxe*, organizado pelas professoras Mariana Fagundes de Oliveira Lacerda e Marian dos Santos Oliveira – as duas com atuação no PROFLETRAS, respectivamente, da UEFS e da UESB –, com apoio financeiro do programa AUXPPG UEFS. O livro se trata do volume 1 da coleção *PROFLETRAS na Prática*, organizada pelas professoras Flávia Aninger de Barros e Mariana Fagundes de Oliveira Lacerda e pelo professor Francisco Fábio Pinheiro Vasconcelos.

Além disso, o NELP disponibiliza produtos educacionais, como o blog Falaê PB (<https://falaepb.wordpress.com/>) – com um Caderno de Aprendizagens de 80 páginas –, desenvolvido como parte da dissertação de Silva (2021)²⁷, orientada pela professora Mariana Fagundes de Oliveira Lacerda.

Está também entre os resultados do projeto NELP na Sala de Aula o mais recente produto: *Tour Virtual pelo Banco CE-DOHS: Coleção Documental Cartas para Vários Destinatários*, trabalho desenvolvido, sob a orientação da professora Mariana Fagundes de Oliveira Lacerda, pela estudante de Graduação Mirian Marques Galindo, vinculada ao Programa de Iniciação Científica da UEFS, com bolsa da FAPESB:

27 Resultados parciais da dissertação de Silva foram publicados em Silva, Lacerda e Tuy Batista (2021).

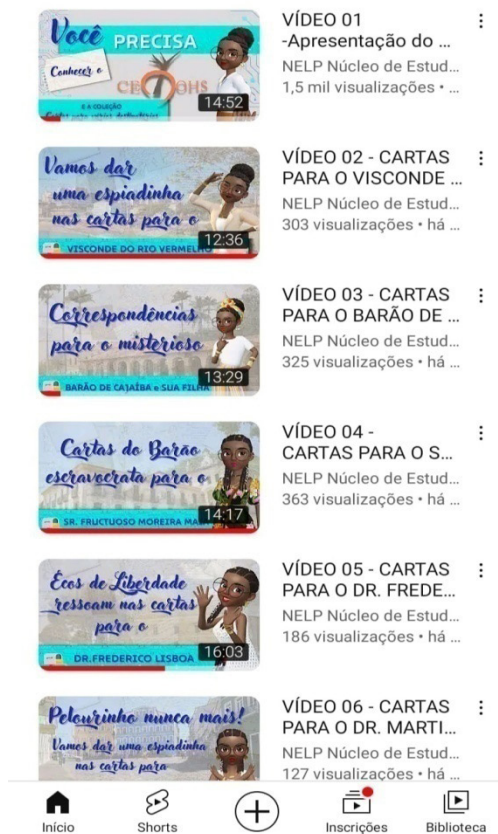


Figura 9: Playlist completa do Tour Virtual pelo Banco CE-DOHS/Cartas para Vários Destinatários.

Fonte: <https://www.youtube.com/@nelpnucleodeestudosemlingu8083/playlists>.

Esse produto – atualmente com 3.895 visualizações no YouTube do NELP – teve ótima repercussão no XXVI Seminário de Iniciação Científica (SEMIC) da UEFS, com muitos elogios, da parte dos parceristas e público ouvinte, também à criação do Avatar do CE-DOHS, Cidinha, a cicerone do *tour*, por Mirian Galindo Marques.

Outro resultado importante, em parceria com o projeto ELiHS, da UNEB, é o jogo ConcordAção, jogo *online* para o ensino da concordância nominal de número (LEMOS, 2023).

Além do projeto NELP na Sala de Aula, há outros dedicados a essa agenda educacional do NELP, como o projeto Corpus Eletrônico de Documentos Históricos do Sertão (CE-DOHS): o banco de dados do NELP/UEFS e sua exploração na interface educacional; Elementos para uma Gramática Digital da Língua Falada no Semiárido Baiano e Plataforma Educacional Digital e Interativa do Núcleo de Estudos de Língua Portuguesa da UEFS, coordenados, de forma colegiada, pela professora Mariana Fagundes de Oliveira Lacerda e pelos doutorandos Janine Araujo da Silva e Wesley da Silva Santos.²⁸ Em desenvolvimento, sob orientação da professora Mariana Fagundes de Oliveira Lacerda, encontram-se duas teses de doutorado, no âmbito do PPGEL UEFS, que terão, entre seus resultados, um produto educacional tecnológico: tese de Janine Silva, Janine Silva (CAPES), *Plataforma Educacional Digital e Interativa do NELP/UEFS: para uma aplicação didático-pedagógica da pesquisa em linguística a partir da elaboração de materiais didáticos com base em dados reais de língua*, e tese de Wesley Santos, *Exploração do banco de textos do NELP/UEFS na interface educacional: elementos para uma Gramática Pedagógica Digital da Língua Falada no Semiárido Baiano*.

Há 20 anos, Vieira (2004, p. 253) afirmou que

O estado atual dos estudos científicos com considerável grau de desenvolvimento, de um lado, e a realidade, muitas vezes infrutífera, do

28 Em desenvolvimento, sob orientação da professora Mariana Fagundes de Oliveira Lacerda, encontram-se duas teses de doutorado, no âmbito do PPGEL UEFS, que terão, entre seus resultados, um produto educacional tecnológico: tese de Janine Silva, Janine Silva (CAPES), *Plataforma Educacional Digital e Interativa do NELP/UEFS: para uma aplicação didático-pedagógica da pesquisa em linguística a partir da elaboração de materiais didáticos com base em dados reais de língua*, e tese de Wesley Santos, *Exploração do banco de textos do NELP/UEFS na interface educacional: elementos para uma Gramática Pedagógica Digital da Língua Falada no Semiárido Baiano*.

ensino de língua portuguesa, de outro, reclamam a aplicação didático-pedagógica da pesquisa em Linguística. (VIEIRA, 2004, p. 253)

Hoje, o estado de coisas não é muito diferente; segundo Lacerda e Silva (2023, p. 105),

o discurso pedagógico, se incorporou o tema da variação linguística – entre outros, relativos a metodologias para o desenvolvimento de atividades de leitura e escrita –, ainda está essa incorporação aquém do que é necessário e desejável. Análises de coleções de livros didáticos evidenciam, por exemplo, que os fenômenos de variação ainda são marginais no ensino de língua materna. (LACERDA; SILVA, 2023, p. 105)

É preciso, como alertam Zilles e Faraco (2015, p. 9), que, como sociedade, discutamos suficientemente, no espaço público, nossa heterogênea realidade linguística e a violência simbólica que a atravessa, de modo a fomentar uma educação linguística plena, a qual se defende também em documentos institucionais²⁹. E, com isso o NELP também assume compromisso e vem empregando esforços, sobretudo a partir de 2017, para uma aplicação didático-pedagógica das pesquisas que realiza, entendendo, como Ilari (2003, p. 111), que um dos principais desafios da Linguística no século XXI é “mudar a cara do ensino da língua materna”

O núcleo tem procurado chegar à sala de aula, em um diálogo com a comunidade docente e discente interna da UEFS e externa, da Educação Básica, seguindo um longo percurso, que começa na formação do banco de textos – cuja prospecção documental depende do conhecimento da realidade sócio-histórica e cuja legitimidade documental, uma vez localizados documentos, depende também dos controles sócio-histórico e paleográfico –, passando pelo estudo linguístico (gramatical e grafo-fonético) até a exploração, na interface educacional, desse material e desses resultados.

29 Conferir Brasil (1998) e Brasil (2016).

É, sem dúvida, um trabalho de fôlego, que envolve uma grande equipe de especialistas, fazendo ciência, de forma coletiva e colaborativa, tendo um programa investigativo claro, porquanto, como acertadamente disse Freitag (2023, p. 1),

a ciência é um trabalho coletivo e colaborativo. Não existe cientista, existe grupo de cientistas que se unem para fazer ciência. A união pode ser no laboratório, mas pode ser também no cafezinho ou no bar. São os encontros, os olhares e as perspectivas diversas que contribuem para o avanço da ciência. (FREITAG, 2023, p. 1)

Os produtos finais do NELP são sempre produtos coletivos, resultantes de uma rotina metodológica bem definida, conforme ilustrado a seguir:

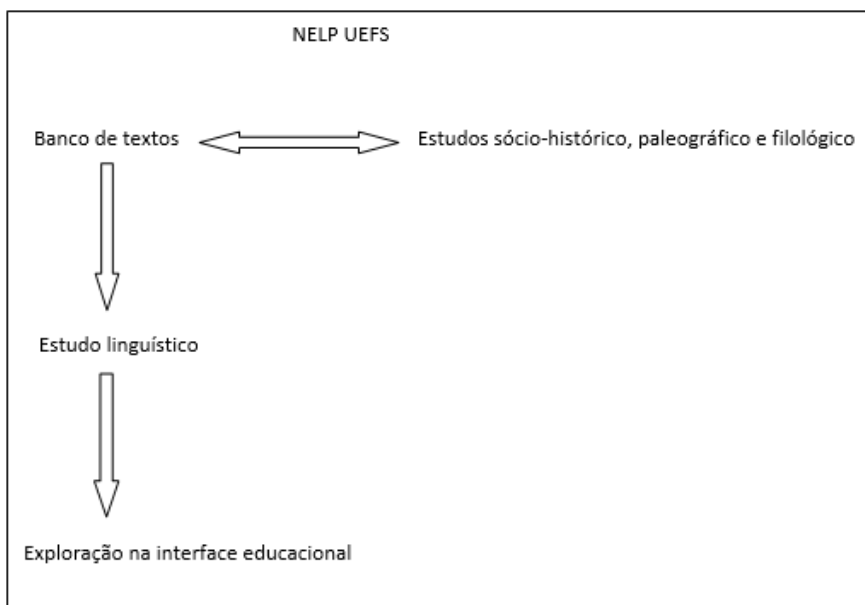


Figura 10: Rotina metodológica do NELP.

Fonte: Elaboração própria.

Palavras finais

História da Língua Portuguesa – cujo objetivo principal é abordar a mudança linguística, tendo em consideração as relações estabelecidas entre a língua em si e a comunidade que a utiliza – é uma disciplina que consta na grade curricular de diversos cursos de Letras, em Instituições de Ensino Superior no Brasil, como na UEFS. Conforme ressaltam Marcotulio *et al* (2018, p. 15), “A formação de um graduando em letras pressupõe a compreensão do dinamismo das línguas, que funcionam como sistemas em constante processo de mudança.” Os alunos da UEFS têm tido, ao longo dos anos, sua formação em Linguística Histórica, história da língua portuguesa e história do PB embasada não apenas nos textos clássicos, que se repetem nos programas das diferentes universidades, mas também nas pesquisas desenvolvidas no NELP, como parte de projetos com equipes interinstitucionais, em rede, dos quais resultam produções importantes, com dados inéditos, que trazem luz à constituição histórica do PB, além de colaborar com discussões teóricas, especialmente sobre tratamento de *corpora*.

O NELP/UEFS se tornou, pela excelência dos trabalhos desenvolvidos, um nome importante no quadro geral de pesquisas em curso no Brasil, sobre a questão complexa da constituição do PB, o qual, como afirmou Mattos e Silva (2004, p. 154), não pode ser tratado como um conjunto homogêneo, unitário, nem numa perspectiva sincrônica, nem numa perspectiva diacrônica. Seja no que diz respeito à constituição de *corpus*, seja no que diz respeito a estudos sócio-históricos e linguísticos, o núcleo se destaca, com trabalhos relevantes, apresentados em eventos acadêmicos, locais, regionais, nacionais e internacionais, e publicados em livros e periódicos, além da disponibilização do *corpus* integral, na rede mundial de computadores, com exploração de dados em diferentes interfaces.

Com o NELP – que recebeu da comunidade acadêmica, em dezembro de 2022, durante a XIX Semana de Letras (SEMALET), na categoria Melhor Núcleo de Pesquisa da UEFS, o prêmio José Jerônimo de Moraes –, a UEFS é reconhecida como um centro de Linguística Histórica no Semiárido baiano,

um dos mais importantes do Brasil, oferecendo à comunidade científica um banco de dados sociolinguísticos para a história do PB, que está entre os mais completos do mundo para a história de uma língua.

A equipe NELP/UEFS vem mergulhando nas “águas profundas” da heterogeneidade complexa do PB, dedicando-se à sua sócio-história passada e presente, (MATTOS E SILVA, 2004, p. 138), colaborando, de forma significativa, com o programa investigativo do PHPB, que é também seu programa investigativo, partindo do Semiárido baiano. E, como não poderia deixar de ser – tendo em vista também o compromisso do núcleo com um ensino de melhor qualidade –, tem feito chegar à Educação Básica, especialmente a escolas de Feira de Santana e cidades circunvizinhas, por meio do projeto NELP na Sala de Aula, a história da língua portuguesa e do português brasileiro, numa abordagem a partir de dados reais, obtidos de *corpora* criteriosamente constituídos.

Fica aqui esta apresentação reflexiva, em síntese, das agendas de trabalho realizadas pelo NELP, podendo servir a experiência do núcleo, como ele se constitui e empreende suas investigações, como modelo a pesquisadores interessados em pesquisas em Linguística Histórica e História do Português Brasileiro.

Ficam aqui, também, para encerrar este texto, nossos agradecimentos à nossa querida UEFS – instituição de que muito nos orgulhamos, sendo para nós uma honra integrar seu corpo docente –, pelo apoio em todos os momentos, às diferentes instituições de fomento que têm dado auxílio ao núcleo, e nossos agradecimentos à professora Rosa Virgínia Mattos e Silva (*in memoriam*), que, em suas aulas na UFBA, ensinou-nos, citando um *Flos Sanctorum* trecentista, que o que salva o homem não é o nome ou a fama, mas a obra que faz (MACHADO FILHO, 2009, fôlio 53vC1)³⁰. Pró Rosa não está

30 “Huum frade preguntou huu monge velho e de gram sanctidade e disse-lhi: - Que é o que salva o homem? O nome ou a fama ou a obra que faz? E o monge velho respondeu i disse: - A obra que faz.”(MACHADO FILHO, 2009, fôlio 53vC1)

mais entre nós, mas sua obra continua a dar bons frutos, e o NELP é um deles, que fica para sempre na história da UEFS e da Linguística brasileira.

Referências

ANDRADE, A. L. de; CARNEIRO, Z. de O. N.; LACERDA, M. F. O. Formas tratamentais em cartas baianas: sujeito e outras funções. In: **Cadernos de Estudos Linguísticos**, 58 (2), 257-276, 2016.

ALMEIDA, N. L. F.; ARAUJO, S. S. F.; PITOMBO, E.; CARNEIRO, Z. O. N.(Org.). **Variação linguística em Feira de Santana-Bahia**. Feira de Santana: Editora UEFS, 2016.

ALMEIDA, N. L. F.; CARNEIRO, Z. O. N. (Org.). **Coleção Amostras da Língua Falada no Semiárido Baiano**. Feira de Santana: Editora UEFS, 2008.

ALMEIDA, N. L. F.; CARNEIRO, Z. O. O NELPRU (Núcleo de Estudos de Língua Portuguesa Rural): apresentação de alguns resultados, de Almeida e Carneiro. In: **Sitientibus**, n. 29, Feira de Santana, 2003. p. 119-132.

BACELAR DO NASCIMENTO, M. F. **O lugar do corpus na investigação linguística**. Disponível em: <<http://www.clul.ul.pt/equipa/berlim-2000-nascimento.pdf>>. Acesso em: 20 abr. 2004.

BARBOSA, A.; LOPES, C. **Críticas, queixumes e bajulações na Imprensa Brasileira do século XIX: cartas de leitores**. Rio de Janeiro: Ed. da UFRJ, 2006.

BARROS, F. A.; LACERDA, M. F. O.; VASCONCELOS, F. F. P. (Coordenação Geral). **Coleção PROFLETRAS em Prática**. 3 volumes. Salvador: Quarteto, 2023.

BORGES NETO, J. **Ensaio de filosofia da linguística**. São Paulo: Parábola, 2004.

BORGES NETO, J. Borges Neto. In: XAVIER, A. C; CORTEZ, S. (Org.). **Conversas com linguistas: virtudes e controvérsias da linguística**. São Paulo: Parábola, 2003. p. 37-50.

BRASIL. Ministério da Educação. **Base Nacional Comum Curricular**. Proposta preliminar. Segunda versão revista. Brasília: MEC, 2016.

BRASIL. Secretaria de Educação Fundamental. **Parâmetros curriculares Nacionais: terceiro e quarto ciclos do Ensino Fundamental: Língua Portuguesa**. Brasília. MEC/SEF, 1998.

BRITO, P. S. J; LACERDA, M. F. O (Org.). Cartas brasileiras: coletânea de fontes para o estudo do português. **Volume 4 (1935-1995): Acervo Cartas Marienses**. Feira de Santana: UEFS Editora, 2022.

BRITO, P.J. S.; LACERDA, M. F. O. A variação da concordância nominal de número em cartas pessoais de Coração de Maria: análise sintagmática e atomística. **Sociodialeto**. V.11, p.1 - 38, 2021.

CAMBRAIA, C. N. **Introdução à crítica textual**. São Paulo: Martins Fontes, 2005.

CARDOSO, L. S.; LACERDA, M. F.; CARNEIRO, Z. O. N. Relações entre a constituição de *corpora* diacrônicos do português brasileiro e os princípios da Ciência Aberta. In: **Revista da Abralin** (no prelo).

CARDOSO, L. S.; CARNEIRO, Z. O. N.; LACERDA, M. F. O. Para um estudo da formação do português brasileiro: descrição, representatividade e potencialidades do corpus colonial do CE-DOHS. In: **LaborHistórico**, v.7, 2021. p.330-355.

CARDOSO, L. S. **A gramática dos pronomes clíticos no Brasil Colônia: o português clássico na história do português brasileiro**. 2020. Dissertação (Mestrado em Estudos Linguísticos) – Programa de Pós-Graduação em Estudos Linguísticos, Universidade Estadual de Feira de Santana, Feira de Santana, 2020.

CARNEIRO, Z. O. N.; TUY BATISTA, P. S.; ALMEIDA, N. L. F. Cartas brasileiras: coletânea de fontes para o estudo do português. **Volume 5 (1920-2000): Acervo da Família Estrela Tuy**. Feira de Santana: UEFS Editora, 2023.

CARNEIRO, Z. O. N.; LACERDA, M. F. O. (Coordenação geral). **Cartas Brasileiras: coletânea de fontes para estudo do português. v. 4 e 5**. Feira de Santana: UEFS Editora, 2023.

CARNEIRO, Z. O. N.; LACERDA, M. F. O. (Org). **CE-DOHS - Corpus Eletrônico de Documentos Históricos do Sertão - Início**. URL: <<http://www5.uefs.br/cedohs/view/home.html>>. Acesso em: 14 fev. 2023.

CARNEIRO, Z. O. N.; LACERDA, M. F. O.; LOBO, T. C. F.; SOUZA, I. L. **Corpora Bahia**. Disponível em: <<https://sites.google.com/site/corporaphbba/?pli=1>>. Acesso em: 14 fev. 2023

CARNEIRO, Z. O. N.; LACERDA, M. F. O. (Coordenação Geral). **Coleção O Sertão por Escrito**: edição filológica dos livros do arquivo do Sobrado do Brejo (Bahia). Feira de Santana: Editora UEFS, 2022, 2 volumes. p. 498.

CARNEIRO, Z. O. N.; LACERDA, M. F. O. ALMEIDA, N. L. F. Indícios sobre a participação dos povos tapuias no contato com o português nos sertões baianos seiscentistas. In: **História Social do Português Brasileiro**: da história social à história linguística. 9 ed. São Paulo: Contexto, 2020, v.9, p. 98-125.

CARNEIRO, Z. O. N.; LACERDA, M. F. O.; ALMEIDA, N. L. F.; SANTIAGO, H. S. O sistema educacional no semiárido baiano oitocentista e notas sobre a escolarização no período republicano feirense e da região sisaleira In: **História Social do Português Brasileiro**: da história social à história linguística. 9 ed. São Paulo: Contexto, 2020, v.9, p. 278-321.

CARNEIRO, Z. O. N.; LACERDA, M. F. de O. Corpus eletrônico de Documentos Históricos do Sertão: etapa 1 (1750-2000). **Revista Binacional Brasil-Argentina: Diálogo entre as ciências**, [S. l.], v. 8, n. 1, p. 205-221, 2019.

CARNEIRO, Z. O. N.; ALMEIDA, N. L. F.; LACERDA, M. F. O.; ARAUJO, S. S. F.; BATISTA, P. S. E. T. Do presente para o passado: os caminhos para a constituição do banco de dados do NELP/UEFS em busca da configuração sócio-histórica da língua portuguesa nos sertões baianos In: **A Fala nordestina: entre a Sociolinguística e a Dialectologia**. 1 ed. Salvador: Blucher, 2016, v.1, p. 106-121.

CARNEIRO, Z. O. N.; OLIVEIRA, M. F. de. (Org.). **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: Editora UEFS, 2012.

CARNEIRO, Z. O. N.; ALMEIDA, N. L. F. Demografia e norma lingüística no semi-árido baiano nos séculos XVIII e XIX: uma introdução. In: Erivaldo Fagundes Neves. (Org.). **Sertões da Bahia: Formação Social, Desenvolvimento Econômico, Evolução Política e Diversidade Cultural**. Salvador: Arcadia, 2011, v. 1, p. 603-628

CARNEIRO, Z. O. N. (Coordenação Geral). **Cartas Brasileiras: coletânea de fontes para estudo do português. Volumes 1, 2 e 3**. Feira de Santana: Editora UEFS, 2011.

CARNEIRO, Z. O. N.; ALMEIDA, N. L. F. Elementos para uma sócio-história da língua portuguesa falada no interior da Bahia. In: Jânia Ramos e Mônica Alkmin. (Org.). **Para a história do português brasileiro - história social**. Belo Horizonte: UFMG, 2008, p. 40-63.

CARNEIRO, Z. O. N.; ALMEIDA, N. L. F. A criação de escolas a partir de critérios demográficos na Bahia do século XIX: uma viagem ao interior. In: LOBO, RIBEIRO, I, CARNEIRO, Z e ALMEIDA, N.. (Org.). **Para a História do Português Brasileiro: novos dados, novas análises**. Salvador: EDFUBA/FAPESB, 2006, v. , p. 649-674.

CASTILHO, A. Projeto **Para a História do Português Brasileiro (PHPBB), Normas de Transcrição de Documentos Manuscritos e Impressos - Edição Semidiplomática**. Disponível em: <<https://sites.google.com/site/corporaphpb/home/normas-de-edicao-do-phpb-2a-versao>>. Acesso em: 14 fev. 2023.

CASTILHO, A.(Org). **História do português brasileiro: o português brasileiro em seu contexto histórico**. São Paulo: Contexto, 2018.

CASTILHO, A. Ataliba de Castilho. In: XAVIER, A. C; CORTEZ, S. (Org.). **Conversas com linguistas: virtudes e controvérsias da linguística**. São Paulo: Parábola, 2003. p. 51-61.

CASTRO, C. **Humanidades Digitais. Estudos Históricos** Rio de Janeiro, vol 33, nº 69, p. 1-2, Janeiro-Abril 2020.

CRANE, G. (et al.). *ePhilology: when the books talk to their readers*. **Blackwell Companion to Digital Literary Studies**. Oxford: Blackwell, 2008.

FARIA, P. F.; KEPLER, F.; PAIXÃO de SOUZA, M. C. **Humanidades Digitais – eDictor**. Disponível em: <<https://humanidadesdigitais.org/edictor/>>. Acesso em: 14 fev. 2023.

FREITAG, R. **Por uma ciência transparente e atual**. Disponível em: <<https://aterraeredonda.com.br/por-uma-ciencia-transparente-e-atual/>>. Acesso em: 5 mar 2023.

FREITAS, M. S; LACERDA, M. F. O.; CARNEIRO, Z. O. N. Formas de tratamento no limiar do século XX: uma análise sociopragmática. **REVISTA DA ANPOLL (ONLINE)**, v.1, p.125 - 155, 2018

GALVES, C.; Andrade, A. L. de; FARIA, P. (2017, December). **Tycho Brahe Parsed Corpus of Historical Portuguese**. Disponível em: . <<http://www.tycho.iel.unicamp.br/~tycho/corpus/texts/psd.zip>>. Acesso em: 14 fev. 2023.

GALVES, C. O Corpus Tycho Brahe: um *corpus* sintaticamente anotado do português histórico. In: **Revista Binacional Brasil-Argentina**, n. 1, 2019. p. 181-204.

GONÇALVES, M. F.; BANZA, A. P. (Org.). **Património textual e humanidades digitais**: da antiga à nova Filologia. Évora: CIDEHUS, 2013.

HOUAISS, A. **O português no Brasil**. Rio de Janeiro: UNIBRADE, 1985.

HERNÁNDEZ-CAMPOY, Juan Manuel; CONDE SILVESTRE, Juan Camilo. **The Handbook of Historical Sociolinguistics**. Oxford: Wiley-Blackwell, 2012.

ILARI, R. Rodolfo Ilari. In: XAVIER, A. C; CORTEZ, S. (Org.). **Conversas com linguistas**: virtudes e controvérsias da linguística. São Paulo: Parábola, 2003. p. 97-112.

LABOV, W. **Padrões sociolinguísticos**. São Paulo: Parábola Editorial, 2008[1972].

LABOV, W. Building on Empirical Foundations. In: Lehmann, W. & Malkiel, Y. (org.). **Perspectives on Historical Linguistics**. Amsterdam/Philadelphia: John Benjamins, 1982. p. 17-92.

LACERDA, M. F. O.; OLIVEIRA, M. S. (Org.). **Ensino de Português**: fonologia, morfologia e sintaxe. v. 1. Salvador: Editora Quarteto, 2023.

LACERDA, M. F. O.; SILVA, J. A. “Cada um fala como quem é”: a morfossintaxe à luz da Sociolinguística Educacional. In: LACERDA, M. F. O.; OLIVEIRA, M. S. (Org.). **Ensino de Português**: fonologia, morfologia e sintaxe. Salvador: Editora Quarteto, 2023, p. 99-134.

LACERDA, M. F.; CARNEIRO, Z. O; SANTIAGO, H. S. (Org.). **Núcleo de Estudos de Língua Portuguesa**. Disponível em: <<https://nelp.uefs.br/>>. Acesso em: 14 fev. 2023.

LACERDA, M. F. O.; SANTOS, E. B. **Livro do Gado (XVIII-XIX)**: edições fac-similar e semidiplomática. Feira de Santana: Editora UEFS, 2022, v.1. p.118.

LACERDA, M. F. O.; ARAUJO, S. S. F.; CARNEIRO, Z. O. N. . Para uma história social linguística nos sertões baianos: elementos para uma proposta de periodização regional. In: ARAÍDE, C.; SILVA, A. P; SILVA, E. C.; ALMEIDA, S. N. J.; RANIERE, T. L. S.; GOMES, V. S. (Org.). **GELNE 40 anos**: vivências teórica e práticas nas pesquisas em linguística e literatura. 1ed.São Paulo: Pá de Palavra, 2018, v. 2, p. 87-106.

LACERDA, M. F. O.; CARNEIRO, Z. O. N.; OLIVEIRA, M. S.; LEMOS, D. M. Formas tratamentais no semiárido baiano: contribuições para uma configuração diatópico-diacrônica do sistema de tratamento do português brasileiro In: **A Fala nordestina**: entre a Sociolinguística e a Dialectologia.1 ed.Salvador: Blucher, 2016, v.1, p. 32-52.

LACERDA, M. F. O.; CARNEIRO, Z. O. N.; SANTIAGO, H. S. Corpus eletrônico de documentos históricos do sertão: as cartas de inábeis. In: **A Cordas Letras**. Vol.17, p.127, 2016.

LACERDA, M. F. O.; CARNEIRO, Z. O. N.; LOPES, B.; ROSÁRIO, T. Edição modernizada dos dois livros manuscritos do Brejo do Campo Seco-Ba (XVIII-XIX): desafios da Nova Filologia. In: **Revista Confluência**, 2023. vol. 64. p. 241-264.

LEMOS, D. M. **ELIHS, um projeto para além do litoral**: da constituição de *corpus* oral de comunidades afro-brasileiras à análise estatística da variação na concordância nominal de número na variedade da comunidade afro-brasileira de Volta do Angico(BA). 2023. Tese de doutorado (Programa de pós-graduação em Estudos Linguísticos). Universidade Estadual de Feira de Santana (UEFS), 2023.

LOBO, T. C. F. Rosa Virgínia Mattos e Silva e a história social linguística do Brasil. **Estudos de Linguística galega**, v. 7, p. 69-82, 2015.

LOBO, T. C. F. Arquivos, acervos e a reconstrução histórica do português brasileiro. In: OLIVEIRA, K.; CUNHA E SOUZA, H. F.; SOLEDADE, J. (Org.). **Do português arcaico ao português brasileiro: outras histórias**. Salvador: EDUFBA, 2009.

LOBO, T. C. F. **Para uma sociolinguística histórica do português no Brasil**. Edição filológica e análise linguística de cartas particulares do Recôncavo da Bahia, século XIX. Volume II. Tese (Doutorado em Filologia e Língua Portuguesa) – Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, 2001

LOPES, C. *et al.* A reorganização do sistema pronominal de 2ª pessoa na história do português brasileiro: a posição de sujeito In: **História do português brasileiro: mudança sintática das classes de palavra: perspectiva funcionalista**. 1 ed. São Paulo: Contexto, 2018, v.1, p. 24-141

LOPES, C. R. S.; MARCOTULIO, L. L.; RUMEU, M. C. B.; LIMA, A. X. Reflexões metodológicas para a análise sociocultural de redatores em *corpora* históricos. In: **Gragoatá**, Niterói, n. 29, 2010.

LOSE, A. D. . Edições de documentos históricos: a quem interessam? A quem se destinam? **Revista da ABRALIN**, v. 16, n. 2, p. 71-86, jan.-abr. 2017.

LOSE, A. D.; PAIXÃO, G.; OLIVEIRA, A. P. S. de; SANCHES, G. A..TELLES, C. M. col. **Dietário do Mosteiro de São Bento da Bahia**: edição diplomática [online]. Salvador: EDUFBA, 2009

LUCCHESI, D. SANTOS, G. **Projeto Vertentes**. Disponível em: <<http://www.vertentes.ufba.br>>. Acesso em: 14 fev. 2013.

LUCCHESI, D. A periodização da história sociolingüística do Brasil. In: **Revista D.E.L.T.A.** 33.2, 2017. Disponível em: <<https://www.scielo.br/j/delta/a/NGxLPBSqNXYNGhFtwqrrwgh/?format=pdf&lang=pt>>. Acesso em: 14 fev. 2023.

MACHADO FILHO, A. V. L. (Ed.). **Um flos sanctorum trecentista em português**. Brasília: Editora da UnB, 2009.

MARCOTULIO, L. L.; LOPES, C. R.; BASTOS, M. J. M.; OLIVEIRA, T. L. **Filologia, história e língua: olhares sobre o português medieval**. São Paulo: Parábola, 2018.

MARQUES, M. G.; LACERDA, M. F. **O.Tour virtual pelo CE-DOHS: Cartas para Vários Destinatários**. Disponível em: <https://youtube.com/playlist?list=PLj5rJKWeYZxNpFsuGwGEmuauc2cJTqocS>. Acesso em: 14 fev. 2023.

MARQUILHAS, R.; HENDRICKX, I. Avanços nas humanidades digitais. In: MARTINS, A. M.; CARRILHO, E. **Manual de Linguística Portuguesa**. MRL Series. De Gruyter, 2016. p. 1-26.

MARTINS, M. A.; ANDRADE, A. L.; MOURA, K. C.; LACERDA, M. F. O. L.; GOMES, V. S. CARNEIRO, Z. O. N. Para um panorama sócio-diacrônico das formas de tratamento na função de sujeito na região Nordeste. **LaborHistórico** 1(1), 26-48, 2015.

MATTOSE SILVA, R. V. O difícil fazer de uma edição crítica de um manuscrito medieval: relato de uma experiência. In: **ESTUDOS LINGÜÍSTICOS**, São Paulo, 42 (2). 2013. p. 669-678.

MATTOS E SILVA, R. V. **Caminhos da linguística histórica: ouvir o inaudível**. São Paulo, SP: Parábola Ed., 2008.

MATTOS E SILVA, R. V. **Ensaio para uma sócio-história do português brasileiro**. São Paulo: Parábola Editorial, 2004a.

MATTOS E SILVA, R. V. **O português são dois**: novas fronteiras, velhos problemas. São Paulo: Parábola, 2004b.

MATTOS E SILVA, R. V. Sobre desencontros e reencontros: Filologia e Lingüística no Brasil no século XX. **Estudos Lingüísticos e Literários**, Salvador, n. 21-22, p. 97-108, 1998.

MATTOS E SILVA, R. V. **Contradições no ensino de português**: a língua que se fala X a língua que se ensina. São Paulo: Contexto, 1996

MATTOS E SILVA, R. V. Estruturas trecentistas: elementos para uma gramática do português arcaico. Lisboa: Imprensa Nacional: Casa da Moeda, 1989.

OLIVEIRA, J. M.; ARAÚJO, S. S. F. O contato entre línguas na constituição da realidade sócio-histórica do português brasileiro: aspectos para um estudo sociolinguístico. In: **A cor das Letras**. v. 19, p. 12, 2018.

PAIXÃO DE SOUZA, M. C. A Filologia Digital em Língua Portuguesa: alguns caminhos. In: **Património Textual e Humanidades Digitais**: Da antiga à nova Filologia. Évora: Publicações do Cidehus, 2013. p. 113-138.

PAIXÃO DE SOUSA, M. C.; KEPLER, F. N.; FARIA, P. E-dictor: Novas perspectivas na codificação e edição de corpora de textos históricos. In: **Anais do VIII Encontro de Linguística de Corpus**, realizado na UERJ, 13 a 14 de novembro de 2009. Rio de Janeiro, 2009. p. 69-105.

PAIXÃO DE SOUSA, M. C.. Lingüística Histórica. In: Claudia Pfeiffer; José Horta Nunes. (Org.). **Introdução às Ciências das Linguagem**: Língua, Sociedade e Conhecimento. 1 ed. Campinas: Pontes, 2006, v. 3, p. 11-48.

PETRUCCI, A. **La ciencia de la escritura. Primera lección de paleografía.** Buenos Aires: Fondo de Cultura Económica de Argentina, 2003.

PPGEL UEFS. **Linguística na Escola.** Disponível em: <<https://www.youtube.com/channel/UCBF6HpKXIZY4QAOMv9dA6ZA>>. Acesso em: 14 fev. 2023.

ROMAINE, Suzanne. **Socio-historical linguistics: its status and methodology.** Cambridge: Cambridge University Press, 1982.

SANTIAGO, H. S.; BRITO, R. C. Pesquisas em Sociolinguística Histórica na Bahia: alguns percursos e possibilidades. In: CARVALHO, C. S. ARAUJO, S. S. F., DIAS, W, C. **A Sociolinguística no Nordeste: diferentes abordagens.** São Paulo: Pá de Palavra, 2023.

SANTIAGO, H. S.; LACERDA, M. F. O.; CARNEIRO, Z. O. N. (Org.). **Cartas em Sisal: estudos morfossintáticos.** Campinas: Editora Pontes, 2023.

SANTIAGO, H. S.; LACERDA, M. F. O.; BRITO, R. C.; CARNEIRO, Z. O. N. CE-DOHS: um banco de dados sociolinguísticos para a história do português brasileiro. In: **LaborHistórico**, vol 7, número especial, Rio de Janeiro. 2021. p. 311-329.

SANTIAGO, H. S. **Um estudo do português popular brasileiro em cartas pessoais de “mãos cândidas” do sertão baiano.** Dissertação (Mestrado em Estudos Linguísticos) – Programa de Pós-graduação em Estudos Linguísticos, Universidade Estadual de Feira de Santana, Feira de Santana, 2012.

SARDINHA, T. B. Linguística de Corpus. In: **Revista D.E.L.T.A.** 23:2. 2007. Disponível em: <<https://www.scielo.br/j/delta/a/yt9hHwfN34PTghT9VdqQJbz/?lang=pt&format=pdf>>. Acesso em: 14 fev. 2023.

SARDINHA, T. B. Linguística de corpus: histórico e problemática. In: **Revista D.E.L.T.A.**, 2000. v. 16. n. 2. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0102-44502000000200005>. Acesso em: 11 mar 2020.

SANTOS, E. B. **O Livro do Gado do Brejo do Campo Seco (Bahia):** edição semidiplomática e descrição de índices grafo-fonéticos. Dissertação (Mestrado em Estudos Linguísticos). Feira de Santana, UEFS, 2019.

SANTOS, E. S. E.; LACERDA, M. F. O.; CARNEIRO, Z. O. N. Sobre formas de tratamento no sertão baiano: a relação entre sujeito e complementos verbais In: **Variação e Mudança na Língua Portuguesa**.1 ed.Campinas: Pontes, 2022, v.1, p. 113-130.

SANTOS, E. S.; LACERDA, M. F. O.; CARNEIRO, Z. O. N. O sistema de tratamento em cartas baianas: uma análise sobre a posição de sujeito. **WORKING PAPERS EM LINGUÍSTICA (ONLINE)**. , v.20, p.109 - 134, 2019a.

SANTOS, E. S.; LACERDA, M. F. O.; CARNEIRO, Z. O. N. Cartas de Sisal (Bahia): Uma Análise Sociopragmática do Sistema de Tratamento In: **Língua e Sociedade: Diferentes Perspectivas, Fim Comum**.1 ed.Salvador: Editora Blucher, 2019b, p. 167-184.

SANTOS, J. V.; BRITO, G. S. Fotografia técnica de documentos para formação de corpora digitais eletrônicos: o método desenvolvido no Lapelinc. *Letras & Letras*, v. 30, n. 2, p. 421-430, 18 dez. 2014.

SHEPHERD, T.; SARDINHA, T. B.; PINTO, M. V. (Org.). **Caminhos da linguística de corpus**. Campinas: Mercado de Letras, 2010.

SILVA, A. J. **O sertão por escrito no Livro de Razão:** um microcosmo sócio-histórico e linguístico da Bahia rural oitocentista. Tese (Doutorado em Língua e Cultura). Salvador, UFBA, 2021.

SILVA, J. A. **Falaê PB**. Disponível em: <https://falaepb.wordpress.com/>. Acesso em: 01 set. 2022.

SILVA, J. A.; LACERDA, M. F. O.; TUY BATISTA, P. S. E. A teoria na prática: a sala de aula como espaço para a pesquisa sociolingüística. In: VILAS BOAS, F. S.; SILVA, O. S. F. **Ensino remoto e formação de professores: construção de novos saberes para os processos educativos**. Campinas: Pontes, 2022. p. 175-204.

SILVA, J. A. **Sociolingüística Educacional e Ensino de Língua Portuguesa: estratégias de pronominalização do objeto direto de terceira pessoa**. Dissertação (Mestrado Profissional em Letras) - Universidade Estadual de Feira de Santana, 2021.

SOUZA, I. L.; PINTO, G. R. P. R.; CARNEIRO, Z. O. N.; FARIA, P.; LACERDA, M. F. O. A ferramenta de busca E-CORP aplicada ao Corpus Eletrônico de Documentos Históricos do Sertão. In: **A Cor das Letras**, v. 19, n. 2, Feira de Santana, 2018. p. 8-21.

TUY BATISTA, P. S. E.; CARNEIRO, Z. de O. N.; LACERDA, M. F. de O. A variação tu/você em relações de solidariedade: análise de uma documentação baiana epistolar do século XX. **CONFLUÊNCIA**. v.2, p.100 - 121, 2017.

TUY BATISTA, P. S.; LACERDA, M. F. O.; CARNEIRO, Z. O. N.; DANTAS Jr., A.; LEAL, I. S. Do lugar de produção/recolha de coleções documentais do CE-DOHS: proposta para uma reconstrução com aplicação de recursos em modelagem 3D. In: **Revista LaborHistórico. Dossiê de Humanidades Digitais**. 2023, v. 9, n. 1, p. 1-27.

VIEIRA, S. R. O ensino da concordância verbal. In: VIEIRA, S. R.; BRANDÃO, S. F. (Org.). **Morfossintaxe e ensino de português: reflexões e propostas**. Rio de Janeiro, Faculdade de Letras da UFRJ, 2004. p. 233-254.

ZILES, A. M. S.; FARACO, C. A. (Org.). Introdução. In: **Pedagogia da variação linguística: língua, diversidade e ensino**. São Paulo: Parábola, 2015. p. 7-18.