

O PORTUGUÊS NO BRASIL COLONIAL: REFLEXÕES SOBRE EDIÇÕES DE MANUSCRITOS HISTÓRICOS BRASILEIROS

PORTUGUESE IN COLONIAL BRAZIL: REFLECTIONS ON EDITIONS OF BRAZILIAN HISTORICAL MANUSCRIPTS

Mariana Fagundes de Oliveira Lacerda | [Lattes](#) | marianafagundes@uefs.br
Universidade Estadual de Feira de Santana

Alicia Duhá Lose | [Lattes](#) | alicialose@gmail.com
Universidade Federal da Bahia

Zenaide de Oliveira Novais Carneiro | [Lattes](#) | zoncarneiro@uefs.br
Universidade Estadual de Feira de Santana

Resumo: Abordamos, neste artigo, o diálogo necessário entre a Paleografia, a Filologia e a Linguística Histórica. Esse diálogo se intensifica no projeto Corpus Eletrônico de Documentos Históricos do Sertão (CE-DOHS), desenvolvido pelo Núcleo de Estudos de Língua Portuguesa (NELP) da Universidade Estadual de Feira de Santana (UEFS), a partir de 2017, quando iniciamos a fase 2 (colonial), atendendo aos objetivos do projeto Guarda-Chuva CE-DOHS: um *corpus* para uma caracterização linguístico-gramatical do português brasileiro - fase colonial e fase pós-colonial, criado em 2010. Neste momento, o CE-DOHS fase 2 já dispõe de coleções documentais em edição semidiplomática, a partir de rigoroso tratamento paleográfico-filológico; além disso, os primeiros estudos de exploração linguístico-gramatical dessa documentação encontram-se realizados. Como sabemos, ainda não é possível atestar a existência do português do Brasil no período colonial brasileiro; mas investigar como se configurava o português no Brasil nesse período pode contribuir para uma reconstrução, por aproximação, da história social linguística do português brasileiro. Nosso objetivo, portanto, no presente texto, é apresentar o ingente trabalho de organização de *corpora* históricos, numa perspectiva interdisciplinar, da maneira como vem sendo realizado no âmbito do projeto CE-DOHS, em parceria com o *Modus Scribendi* – Grupo de Pesquisas Paleográficas, Filológicas e Históricas, da Universidade Federal da Bahia (UFBA); dessa organização dependem descrições linguísticas de base empírica, as quais promovem a discussão de questões teóricas solidamente fundamentadas.

Palavras-Chave: Português no Brasil; *Corpus* colonial; Tratamento paleográfico-filológico.

Abstract: In this article, we address the necessary dialogue between Paleography, Philology and Historical Linguistics, a dialogue that intensifies in the Electronic Corpus of Historical Documents of the Sertão (CE-DOHS) project, of the Portuguese Language Studies Center (NELP), of State University of Feira de Santana (UEFS), from 2017, when we started phase 2 (colonial), meeting the objectives of the Umbrella CE-DOHS project: a corpus for a linguistic-grammatical characterization of Brazilian Portuguese - colonial phase and post-colonial phase, created in 2010. At this moment, CE-DOHS phase 2 already has documentary collections in a semi-diplomatic edition, based on rigorous paleographic-philological treatment, as well as the first linguistic-grammatical exploration studies of this documentation found carried out. As we know, it is not yet possible to attest to the existence of Brazilian Portuguese in the Brazilian colonial period; but investigating how Portuguese was configured in Brazil, during this period, can contribute to an approximation reconstruction of the linguistic social history of Brazilian Portuguese. Our objective, therefore, in this text, is to present the enormous work of organizing historical corpora, from an interdisciplinary perspective, as it has been carried out within the scope of the CE-DOHS project, in partnership with Modus Scribendi – Paleographic Research Group, Philological and Historical, from the Federal University of Bahia (UFBA); this organization depends on empirically based linguistic descriptions, which promote the discussion of solidly based theoretical issues.

Keywords: Portuguese in Brazil; Colonial corpus; Paleographic-philological treatment.

1 Introdução ou sobre o trabalho para muitas mãos...

Desde 1997, uma grande equipe de pesquisadores – que só tem aumentado ao longo dos anos dedica-se, no âmbito do Projeto Nacional para a História do Português Brasileiro (PHPB), à reconstrução, por aproximação, da história social linguística do PB (Mattos e Silva, 2004), atuando em diferentes agendas de trabalho (Castilho, 2024). Entre elas, a agenda de *corpus* diacrônico, com a qual o projeto Corpus Eletrônico de Documentos Históricos do Sertão (CE-DOHS)¹⁻² – do Núcleo de Estudos de Língua Portuguesa (NELP)³ da Universidade Estadual de Feira de Santana (UEFS) – tem contribuído de forma bastante significativa.

¹ Disponível em: <http://www.uefs.br/cedohs/>. Link de acesso ao Grupo no CNPq: <http://dgp.cnpq.br/dgp/espelhogrupo/481138>.

² 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.

³ Disponível em: <https://nelp.uefs.br/>.

Como afirmou Mattos e Silva (2004, p. 61), a busca pela reconstrução da história do português brasileiro “será trabalho para muitas mãos, durante muito tempo”. E esse trabalho inclui, fundamentalmente, a constituição de *corpus*, em que buscamos os dados que iluminam a linguística teórica. A mão do paleógrafo e a mão do filólogo, juntas à mão do historiador das línguas, garantem a autenticidade dos materiais, dos *corpora*, cujo uso “permite a realização de descrições linguísticas de base empírica e promove, com isso, a discussão de questões teóricas solidamente fundamentadas” (Bacelar do Nascimento, 2004, p. 1).

Em publicação de 2008a, Mattos e Silva ressalta 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, 2008a, p. 15).

Como vemos, esse cruzamento de olhares – diplomático, paleográfico, filológico, sócio-histórico e linguístico – no trabalho de constituição de *corpus* em Linguística Histórica *Stricto Sensu* (Mattos e Silva, 2008) é essencial, condição *sine qua non* para o estudo da mudança linguística no tempo real. A história das línguas, como ressalta Mattos e Silva (2008b, p. 52), citando o historiador José Mattoso, é feita, similar à História, da “incomensurável mistura de elementos de toda espécie”.

O historiador medieval José Mattoso, no seu livro “A escrita da história: teoria e métodos” diz: “Prefiro... tentar descobrir as harmonias resultantes dessa fantástica sinfonia que é a História, feita da incomensurável mistura de elementos de toda espécie, tão dispersos e contraditórios com a própria vida, mas de cuja rede infinitamente complexa e fascinante procuram os eixos, os encontros e desencontros” (1988: p. 10). O que disse o historiador para a História aplica-se, a meu ver, à história da(s) língua(s) (Mattos e Silva, 2008b, p. 52).

Pisando nesse terreno complexo e fascinante da história das línguas, abordamos, neste artigo, um diálogo necessário entre a Paleografia, a Filologia e a Linguística Histórica. Esse diálogo, que se intensifica no projeto CE-DOHS/NELP/UEFS a partir de 2017, quando iniciamos a fase 2 (colonial), atende aos objetivos do projeto Guarda-Chuva CE-DOHS: um *corpus* para uma caracterização linguístico-gramatical do portu-

guês brasileiro - fase colonial e fase pós-colonial, criado em 2010 por Zenaide de Oliveira Novais Carneiro e Mariana Fagundes de Oliveira Lacerda. Abordamos, ainda, a importância da pesquisa em rede, como a que vem sendo desenvolvida no âmbito do PHPB, há 28 anos, tendo como coordenador Ataliba Teixeira de Castilho; e no âmbito do NELP, criado em 1997 por Zenaide de Oliveira Novais Carneiro e Norma Lucia Fernandes de Almeida e coordenado por Mariana Fagundes de Oliveira Lacerda⁴, desde 2017.

Nosso objetivo, portanto, no presente texto, é apresentar o ingente trabalho de organização de *corpora* históricos, numa perspectiva interdisciplinar, da maneira como vem sendo realizado no âmbito do projeto CE-DOHS, em parceria com o *Modus Scribendi* – Grupo de Pesquisas Paleográficas, Filológicas e Históricas da Universidade Federal da Bahia (UFBA), coordenado por Alícia Duhá Lose e Livia Borges Souza Magalhães.

2 OPHPB e a enormidade do empreendimento

Em janeiro de 2024, por ocasião do XX Congresso Internacional da Associação Brasileira de Linguística e Filologia da América Latina (ALFAL), no âmbito do Projeto 3 - História do Português Brasileiro, da Europa à América, realizado na Faculdade de Humanidades e Arte da Universidade de Concepción, no Chile, o professor Ataliba Teixeira de Castilho (USP/UNICAMP) proferiu a Conferência de encerramento, intitulada *Historiando a Língua Portuguesa no Novo Mundo*. No texto, ele afirma:

A Linguística Histórica perdeu espaço entre nós durante o período mais forte de atuação do Estruturalismo e do Gerativismo, movimentos basicamente a-históricos. É verdade, entretanto, que, mesmo no interior dessas teorias, algumas vozes anunciavam a presença da “velha senhora”, como gosto de qualificar a Linguística Histórica. (...). No Brasil, foi profética a voz de Tarallo (1984), que proclamou o renascimento da “fênix”, excelente metáfora para uma ciência que retorna sempre. Nos anos 1980, sem que tivesse havido uma combinação prévia, três respeitadas linguistas fizeram renascer a Linguística Histórica no domínio da língua portuguesa: a) Clarinda Maia, da Universidade de Coimbra; b) Rosa Virgínia Mattos e Silva, da Universidade Federal da Bahia; c) Mary Kato, da Universidade Estadual de Campinas. O quadro desenhado acima me motivou a propor a criação de um projeto coletivo de pesquisas, tendo

⁴ Conferir linha do tempo do NELP em: <https://nelp.uefs.br/o-nelp/historico/>. O NELP/UEFS completou 25 anos em 2023, sendo considerado um centro de Linguística Histórica no Semiárido baiano. O PHPB, desde sua criação, em 1997, tem contado com a equipe de pesquisadores do NELP e recebido contribuições significativas às agendas de trabalho. Conferir Lacerda (2024), NELP 25 anos <https://nelp.uefs.br/nelp-25-anos/> e Livros Publicados/Coleções de *Corpora* e Coletâneas Linguísticas <https://nelp.uefs.br/livros-publicados/>.

por objetivo, inicialmente, historiar o Português de São Paulo (Castilho, 1998) (Castilho, 2024, p. 2)⁵.

O Projeto para a História do Português Paulista (PHPP) realizou o seu I Seminário em 1997, e, nessa oportunidade, por iniciativa dos participantes, o projeto regional assumiu uma dimensão nacional, criando-se o Projeto para a História do Português Brasileiro (PHPB). Castilho (2024, p.3) destaca que “O ‘para’ foi uma sugestão oportuna de Mattos e Silva, que nos advertia sempre sobre a enormidade do empreendimento”. Um projeto nacional, com 26 anos de existência, que se acha atualmente entre os maiores projetos em rede de pesquisa da Linguística brasileira.

Sem dúvida, um empreendimento enorme a reconstrução, por aproximação, da história social linguística do PB. No texto supracitado de 2024, Castilho refere-se às seguintes agendas de trabalho do PHPB:

Desde sua criação, o PHPB foi se consolidando como um projeto nacional, desenvolvendo as seguintes agendas de trabalho: a) Corpus diacrônico; b) História social; c) Mudança gramatical; d) Linguística diacrônica do texto; e) Tradições discursivas; f) Semântica diacrônica; g) História do léxico. Nos primeiros 10 anos, as atividades se concentraram na organização do *corpus* diacrônico e na mudança gramatical, a partir de perspectivas teóricas mais sócio-funcionais ou formais. Com a ampliação das equipes e a necessidade de consolidar os resultados do projeto, novas temáticas foram incorporadas ao Projeto (Castilho, 2024, p. 4).

A agenda de *corpus* diacrônico, desde sempre um grande desafio, trata-se de uma das principais agendas do projeto, tendo a Comissão de *Corpus*, em 1º de junho de 2010, definido “o conjunto do *Corpus* Mínimo Comum Diacrônico (séculos XVIII, XIX e XX)”. Segundo Castilho (2024, p. 7)⁶, “[a] opção foi compor três conjuntos de textos divididos da seguinte maneira: *Corpus* comum mínimo – manuscritos; *Corpus* comum mínimo – impressos; *Corpus* comum diferencial”.

A Bahia, como abordamos nas Seções 3, 4 e 5, adiante, está entre as equipes regionais do PHPB, destacando-se no trabalho de constituição de *corpus*, entre outras agendas. O Quadro 1 apresenta os membros das equipes regionais do projeto.

⁵ O autor esclarece a menção às três respeitadas linguistas: “a) Clarinda Maia, da Universidade de Coimbra, com seus estudos sobre o Galego-Português: Maia (1986). Ver também Maia (2022). b) Rosa Virgínia Mattos e Silva, da Universidade Federal da Bahia, com seu estudo sobre o Português trecentista, a que se seguiu o *Programa de História do Português* (doravante, PROHPOR): Mattos e Silva (1989). c) Mary Kato, na Universidade Estadual de Campinas, juntamente com Fernando Tarallo, quando focalizaram o Português Brasileiro em trabalho apresentado em um congresso internacional: Kato; Tarallo (1988). Depois, eles deram início a um projeto de história do português brasileiro, de que surgiram várias teses e ensaios, já publicados: Roberts, Kato (1993); Galves, Kato e Roberts (2019)” (Castilho, 2024, p. 2).

⁶ Plataforma de *Corpora* PHPB disponível em: <https://sites.google.com/site/corporaphpb/home>.

Quadro 1: Equipes regionais do PHPB.

Alagoas	Thiago Trindade (Universidade Federal de Alagoas).
Bahia	Pedro Daniel dos Santos Souza (Universidade do Estado da Bahia), em substituição a Tânia Lobo, que substituiu Rosa Virgínia Mattos e Silva.
Ceará	Aurea Suely Zavam (Universidade Federal do Ceará).
Mato Grosso	Elias Alves de Andrade (Universidade Federal do Mato Grosso).
Minas Gerais	Jânia Ramos (Universidade Federal de Minas Gerais).
Paraíba	Roseane Nicolau, em substituição a Camilo Rosa (Universidade Federal da Paraíba).
Paraná	Vanderci Aguilera (Universidade Estadual de Londrina).
Pernambuco	Cléber Alves de Ataíde, em substituição a Valéria Gomes (Universidade Federal Rural de Pernambuco), que substituiu Marlos de Barros Pessoa (Universidade Federal de Pernambuco).
Rio de Janeiro	Maria Eugênia Duarte, em substituição a Dinah M. I. Callou (Universidade Federal do Rio de Janeiro).
Rio Grande do Norte	Alessandra Castilho da Costa, em substituição a Marco Antonio Martins (Universidade Federal do Rio Grande do Norte).
Santa Catarina	Izete L. Coelho, em substituição a Gilvan Muller de Oliveira (Universidade Federal de Santa Catarina).
São Paulo	Manoel Mourivaldo Santiago Almeida (Universidade de São Paulo), em substituição a Clélia C. S. Jubran (Universidade Estadual Paulista / São José do Rio Preto), que substituiu Ataliba T. de Castilho (Universidade de São Paulo).
Sergipe	Sandro Márcio Drumond Alves Marengo, em substituição a Renata Ferreira Costa (Universidade Federal de Sergipe).

Fonte: Elaborado pelas autoras, com base em Castilho (2024, p. 5).

As equipes regionais do PHPB se reúnem em seminários nacionais desde 1998; nesse ano, o seminário ocorreu em São Paulo, e o último, em 2019, ocorreu em Sergipe, em comemoração aos 20 anos do projeto.

Os seminários nacionais e regionais do PHPB produziram atas⁷. Por exemplo, no volume IV (Duarte; Callou, 2002), sobre *corpus*, têm destaque, entre outros, os seguintes temas:

- a) O controle sócio-histórico da documentação, tanto quanto possível;
- b) A identificação do perfil biográfico de quem escreve;

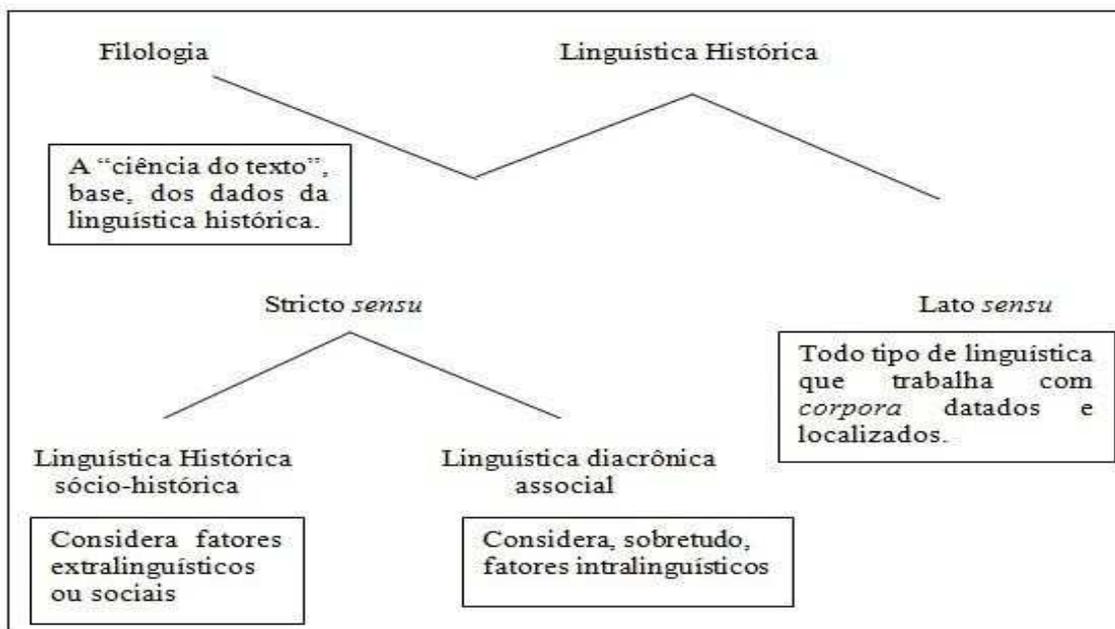
⁷ Produziram-se, até o presente, 16 atas, além dos 10 volumes do PHPB, publicados pela editora Contexto.

- c) O produtor do texto (autenticidade do material e grau de domínio da técnica de escrever de cada autor/*escriptor*);
- d) O contexto de produção do texto (as relações sociais que cercam o ato de escrever na época em foco).

Os controles sócio-histórico, paleográfico e filológico, no trabalho de constituição de *corpus* diacrônico no âmbito do PHPB, sempre foram tidos como condição *sine qua non*. Questões sobre a autenticidade dos documentos, sobre sua leitura, sobre sua transcrição, entre outras, estavam, desde o princípio, na *mesa* do PHPB.

Demonstrando a dependência da Linguística Histórica *Stricto Sensu* em relação à Filologia, disciplinas parceiras na coletividade científica⁸, Mattos e Silva (2008a) constrói um esquema apresentado na Figura 1.

Figura 1: Parceria entre a Linguística Histórica *Stricto Sensu* e a Filologia



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

Sem dúvida, olhares entrelaçados, tendo em vista edições que satisfaçam as necessidades de uma Linguística Histórica empírica (Maia, 2012, 538).

⁸ Conferir a discussão sobre a parceria entre Linguística Histórica e Filologia feita por Lose (2017) e Lose e Souza (2022).

3 A agenda de *corpus* diacrônico na Bahia: como formigas obreiras

O PHPB enfrentou grandes desafios no trabalho de constituição de *corpus* diacrônico, especialmente na primeira década. Segundo Lobo,

[a] segunda metade da década de 1990 inaugura, nos estudos histórico-diacrônicos do português brasileiro, uma fase que se caracterizará pela laboriosa tarefa de construção de uma filologia de textos escritos no Brasil, ponto de partida incontornável para a descrição e análise das mudanças linguísticas que foram configurando o chamado português brasileiro. A inexistência de uma tradição filológica relativa ao período compreendido entre os séculos XVII e o XIX era, então, observável não apenas no Brasil, mas também em Portugal (Lobo, 2009, p. 307-308).

Dado esse cenário, Lobo (2009), que foi uma das coordenadoras do PHPB-Bahia, esclarece que, na primeira década, a constituição de *corpus* diacrônico na Bahia passou por algumas fases: a fase ingênua; a fase da profissionalização; e a fase dos profissionais.

A fase ingênua é caracterizada pela edição de textos escritos no Brasil, sem, ainda, o conhecimento necessário sobre os arquivos, sobre os fundos documentais, sobre a caracterização sociolinguística dos informantes (Lobo, 2009).

A fase da profissionalização remete ao período em que

não bastava editar textos escritos no Brasil e, desavisadamente, tomá-los como representativos de um português brasileiro, já que, no Brasil do passado, a pena esteve nas mãos de portugueses, “brancos” brasileiros, africanos, índios e imigrantes – todos, com maior ou menor perícia, escrevendo em português. Dentre as diversas questões a enfrentar, uma sobressaía-se: era preciso identificar, o mais precisamente possível, “quem escreveu” (Lobo, 2009, p. 312)

Por último, a fase dos profissionais é entendida por Lobo (2009, p. 314) desta forma: “Aqui se alcança um grau de maturidade no rastreamento de fontes primárias as mais diversas. Despontam os profissionais”.

Como ressalta Mattos e Silva (2008b, p. 39-40), a constituição de *corpus* tem início na prospecção documental, um trabalho de “formigas obreiras” pelos arquivos, em vários pontos do Brasil:

A Filologia se fez presente no PHPB: no Seminário de Campos do Jordão, um grupo, coordenado por Heitor Megale, estabeleceu as normas para as edições a serem utilizadas no *corpus* documental que seria a base das análises linguísticas. Começou-se com textos impressos no século XIX, sobretudo os jornais, que, no Brasil, se multiplicaram naquele século, uma vez que, só

a partir de 1808, foi permitida pela metrópole a imprensa no Brasil. Dessa pesquisa em jornais, publicou-se o livro *E os preços eram commodos... Anúncios de jornais brasileiros. Século XIX*, sob a responsabilidade de Marymárcia Guedes e Rosane de Andrade Berlinck, corria o ano 2000. Depois foi-se para os Arquivos Brasileiros, públicos ou privados. A equipe da Bahia publicou, em 2001, *Cartas Baianas Trecentistas*, sob a responsabilidade da pesquisadora Tânia Lobo e com a colaboração dos bolsistas de Iniciação Científica – Klebson Oliveira, Permínio Ferreira e Uilton Oliveira. Outros *corpora* vêm sendo constituídos, como base, principalmente, de dissertações e teses, como, por exemplo – *As cartas do Século XIX* para as freiras do Desterro, base da tese de Tânia Lobo (inédita), em que a autora distingue as cartas de brasileiros das dos portugueses; as *Cartas* de Afrânio Gonçalves de comerciantes portugueses no Brasil, que também editou outros documentos para confrontar com as *Cartas*; as atas e outros documentos da Sociedade Protetora dos Desvalidos, editadas por Klebson Oliveira, no seu Mestrado, e ampliadas, em número, na sua tese de Doutorado, e assim por diante. Verificou-se, no correr da pesquisa, que era o século XIX o mais fértil nos Arquivos Brasileiros, diminuindo a incidência de documentos quando se recua no tempo. Como formigas obreiras, têm-se pesquisado Arquivos em vários pontos do Brasil (Mattos e Silva, 2008b, p. 39-40).

Novos olhares, especialmente da Paleografia, sobre os *corpora* constituídos no âmbito do PHPB Bahia⁹ ao longo dos anos, bem como no âmbito de outras equipes regionais do PHPB nacional, poderão lançar luz sobre aspectos específicos dos documentos, numa rede de pesquisa colaborativa, produtiva e que prima pelo rigor e pela honestidade científicos. Como bem destacou Maia (2012, p. 538),

[s]ob 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).

Portanto, é sempre necessário, e desejável, que os materiais para exploração de dados linguísticos, na Linguística Histórica, recebam os tratamentos diplomático, paleográfico e filológico com o devido rigor científico, além de ser realizado, tanto quanto possível, controle sócio-histórico rígido, de modo que os estudos linguísticos realizados com esses *corpora* tenham confiabilidade. Isso porque a confiabilidade dos dados depende da autenticidade dos materiais.

A promoção dessa rede de pesquisa, entre historiadores das línguas, paleógrafos e

⁹ O Grupo de Pesquisa *Modus Scribendi* – Grupo de Pesquisas Paleográficas, Filológicas e Históricas (CNPq-UFBA), por exemplo, está revisitando a documentação da Sociedade Protetora dos Desvalidos (SPD), com que trabalhou Oliveira (2006), trazendo novos documentos e novos elementos para discussão.

filólogos, em um diálogo respeitoso, considerando os diferentes contextos de produção de pesquisa, espaços e tempos, favorece o melhor uso, em Linguística Histórica, dos dados disponíveis. O trabalho interdisciplinar, coletivo, que, nas palavras de Mattos e Silva¹⁰, “desenvolve a solidariedade exigente e permite o intercâmbio fraterno entre os saberes de cada um”, vai formando pesquisadores com maior grau de maturidade em qualquer campo do conhecimento.

4 CE-DOHS: fontes para a história do português brasileiro

Faz parte do NELP/UEFS, da Bahia, como dissemos na introdução deste texto, o projeto guarda-chuva intitulado CE-DOHS: um *corpus* para uma caracterização linguístico-gramatical do português brasileiro - fase colonial e fase pós-colonial, criado em 2010 e coordenado por Zenaide de Oliveira Novais Carneiro e Mariana Fagundes de Oliveira Lacerda:

Trata-se de um projeto guarda-chuva – vinculado ao Núcleo de Estudos de Língua Portuguesa (NELP), da UEFS, com 25 anos de existência – que se dedica à constituição de *corpora* históricos de português brasileiro (fase colonial e fase pós-colonial), a partir de cuidadosa prospecção documental e estabelecendo rigorosos controles sócio-histórico, paleográfico e filológico, tendo em vista colaborar com os estudos sobre a história social linguística do PB (Mattos e Silva, 2004), em parceria com o Projeto Nacional para a História do Português Brasileiro (PHPB) e com o Programa para a História da Língua Portuguesa (PROHPOR). Este projeto, que tem início em 2010, trata-se da continuação do projeto, intitulado Contribuições para a Constituição de um Banco de Textos e de um Banco de Dados para o Estudo da História do Português do Brasil, do Século XVII ao XX e iniciado no NELP/UEFS, em 1997, sob a coordenação de Ilza Ribeiro. Vê-se, daí, portanto, a longa trajetória que tem o NELP/UEFS no que diz respeito a estudos sobre o PB nas fases colonial e pós-colonial, firmando-se como uma das principais instituições responsáveis por essas agendas no Brasil, tendo reconhecimento internacional. (...) São significativos os resultados alcançados até o presente, especialmente no âmbito de dissertações de mestrado e teses de doutorado. Consistem em análises sobretudo de natureza qualitativa (especialmente para o período colonial, haja vista a raridade dos documentos e os dados escassos), as quais dependem crucialmente de interpretações de base teórica, refinadas pela Paleografia (Carneiro; Lacerda, 2024, Início).

Segundo Ribeiro (1997), no seu projeto – motivador da criação do NELP –, intitulado Contribuições para a Constituição de um Banco de Textos e de um Banco de Dados para o Estudo da História do Português do Brasil, do Século XVII ao XX, com os direcionamentos propostos e assumidos pelo PHPB – entre os quais a constituição de um *corpus* diacrônico do PB –,

¹⁰ Citação de Mattos e Silva no *menu* Histórico, do site do PROHPOR UFBA.

está bem pensada e planejada uma reconstrução dos processos lingüísticos e socioculturais condicionadores da formação do português no Brasil, ou português brasileiro, a partir de uma organização sistemática e de estudos lingüísticos de diferentes documentos (estilos diversos) escritos no Brasil, em diversas regiões e em diferentes épocas de sua história. Houaiss (1995: 31) diz que “não preenchemos ainda os requisitos da pesquisa e conhecimento com que se possa elaborar uma história da língua portuguesa no Brasil”. Certamente essas pesquisas mudarão o estado da questão (Ribeiro, 1997, p. 3-4).

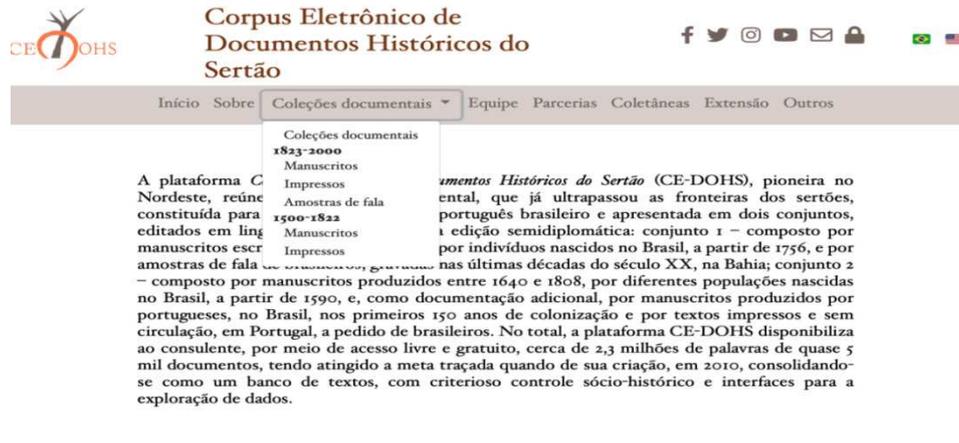
De fato, muito já foi feito, ao longo de mais de duas décadas, pela equipe do PHPB nacional, e importantes contribuições lançaram luz à história do PB.

Atendendo à agenda de *corpus* diacrônico, o PHPB, ao longo de sua história, reuniu documentação sobretudo do período pós-colonial, dos séculos XIX e XX. O CE-DOHS/NELP/UEFS, a partir de 2017 – já disponibilizando rico material da fase 1 (pós-colonial) –, começou a investir na formação de *corpus* ilustrativo do período colonial, intensificando a parceria com a Paleografia e com a Filologia para o tratamento dos textos de épocas ainda mais recuadas¹¹; conferir figura 2. Destaca-se a parceria com o Grupo de

¹¹ Esse estreitamento no diálogo da Linguística Histórica com a Paleografia e a Filologia tem sido promovido pelo CE-DOHS e pelo *Modus Scribendi* também em eventos acadêmicos, como: (i) VI Congresso Internacional de Linguística Histórica (Faculdade de Letras da Universidade de Coimbra, 2023) – “Edições de manuscritos históricos brasileiros: a série Documentos para a história linguística do Brasil colônia”, por Alícia Duhá Lose, Mariana Fagundes de Oliveira Lacerda e Zenaide de Oliveira Novais Carneiro; (ii) XX Congresso Internacional da Associação Brasileira de Linguística e Filologia da América Latina - Projeto 3 da ALFAL, História do Português Brasileiro, da Europa à América (Faculdade de Humanidades e Arte da Universidade de Concepción, Chile, 2024) – “O português no Brasil colonial: reflexões sobre as edições de manuscritos históricos brasileiros”, por Alícia Duhá Lose, Zenaide de Oliveira Novais Carneiro, Mariana Fagundes de Oliveira Lacerda e Lara da Silva Cardoso; (iii) XI Seminário de Estudos Filológicos: Filologia e Ensino - Perspectivas Críticas e Editoriais (Universidade Estadual de Feira de Santana, Feira de Santana, 2024) – “O CE-DOHS na rede de pesquisa PHPB: recuando ao período colonial”, por Mariana Fagundes de Oliveira Lacerda, Zenaide de Oliveira Novais Carneiro e Alícia Duhá Lose; (iv) Simpósio da Contemporaneidade à Antiguidade: temas em debate (Centro de Estudos Clássicos e Humanísticos da Universidade de Coimbra) – “Reflexões sobre edições de manuscritos históricos brasileiros: o cruzamento de olhares paleográfico e linguístico”, por Alícia Duhá Lose, Mariana Fagundes de Oliveira Lacerda e Zenaide de Oliveira Novais Carneiro; (v) IX Simpósio Mundial de Língua Portuguesa (Universidade da Madeira, Ilha da Madeira, Portugal, 2024) – “Fontes linguísticas para estudos do português no Brasil colonial: o exemplo do Testamento de André Vidal de Negreiros (séc. XVII), por Alícia Duhá Lose, Mariana Fagundes de Oliveira Lacerda, Zenaide de Oliveira Novais Carneiro e Victória Santana da Silva Araújo; e (vi) 30ª Jornada do GELNE, em dezembro de 2024 (Universidade Estadual de Feira de Santana, Feira de Santana) – “O trabalho de organização de *corpora* históricos, a pesquisa empírica e a periodização da história sociolinguística do Brasil, por Mariana Fagundes de Oliveira Lacerda, Zenaide de Oliveira Novais Carneiro e Alícia Duhá Lose. Em outubro de 2025, no evento SociolinguísticaS, na Universidade Federal da Paraíba (UFPB), apresentaremos o trabalho, intitulado *Plataforma CE-DOHS: língua, história e tecnologia*, na Mesa sobre *Repositório de dados de e para pesquisa*, sob a coordenação da professora Márcia dos Santos Machado Vieira, da Universidade Federal do Rio de Janeiro (UFRJ).

Pesquisa *Modus Scribendi* – Grupo de Pesquisas Paleográficas, Filológicas e Históricas, da Universidade Federal da Bahia (UFBA), anteriormente mencionado.

Figura 2: Página das Coleções Documentais do CE-DOHS



Fonte: Disponível em: http://www.uefs.br/cedohs/view/colecoes_documentais.html.

A fase 2 (colonial) teve início em 2017 com a execução da pesquisa intitulada *O Livro do Gado do Brejo do Campo Seco (Bahia): edição semidiplomática e descrição de índices grafo-fonéticos*, dissertação de Elaine Brandão Santos, desenvolvida no Programa de Pós-Graduação em Estudos Linguísticos (PPGEL) da UEFS. Trata-se de um documento do final do período colonial, da fazenda do Brejo do Campo Seco (BA).

Segundo Santos Filho, o *Livro do Gado* e o *Livro de Razão* (este editado por Silva [2021]), em tese intitulada *O sertão por escrito no Livro de Razão: um microcosmo sócio-histórico e linguístico da Bahia rural oitocentista*, constituem “[v]elhos memoriais de 200 anos, manuseados e ensebados quotidianamente pelas mãos de três gerações, guardados depois pela quarta e pelas seguintes, metamorfosearam-se hoje em preciosas relíquias, preciosas testemunhas de remoto passado” (2012, p. 110).

Esses dois livros de fazenda – o *Livro do Gado*, com 57 folhas, escrito de 1755 a 1832, e o *Livro de Razão*, com 195 folhas, escrito de 1795 a 1838 – tiveram suas edições fac-similar e semidiplomática¹² publicadas em 2022 na coleção *O Sertão por Escrito: edição filológica dos livros do arquivo do Sobrado do Brejo (Bahia)*, coordenada por Zenaide de Oliveira Novais Carneiro e Mariana Fagundes de Oliveira Lacerda¹³.

¹² O CE-DOHS/NELP/UEFS usa as normas de transcrição do PHPB (Castilho, 2020) e do Centro de Estudos, Pesquisa e Documentação Paleográfica do Memória e Arte (CEPEDOP).

¹³ A orelha da coleção foi escrita pela paleógrafa e filóloga Alicia Duhá Lose (UFBA/UEFS).

Figura 3: Capas dos dois volumes da Coleção *O Sertão por Escrito*



Fonte: Disponível em: <http://www.uefs.br/cedohs/view/coletaneas.html/>.

Neste momento, como resultado de parceria de pesquisa entre o CE-DOHS/NELP/UEFS e o Grupo *Modus Scribendi*/UFBA – a partir de um cruzamento de olhares paleográfico, filológico, sócio-histórico e linguístico –, encontra-se, no prelo, a *série Documentos para a História Linguística do Brasil Colônia*, sob a coordenação geral de Zenaide de Oliveira Novais Carneiro, Mariana Fagundes de Oliveira Lacerda e Alícia Duhá Lose, já anunciada na página *online* do CE-DOHS¹⁴, com publicação pela UEFS Editora, inicialmente com 16 volumes¹⁵. São eles:

Quadro 2: Volumes a sair na série Documentos para a História Linguística do Brasil Colônia, pela UEFS Editora

Volume 1	Navarro, Eduardo; Lose, Alícia Duhá; Carneiro, Bruna Trindade Gomes; Lacerda, Mariana Fagundes de Oliveira; Carneiro, Zenaide de Oliveira Novais. Cartas dos Índios Camarões escritas em Tupi Antigo (1645).
Volume 2	Araújo, Victória da Silva Santana; Lacerda, Mariana Fagundes de Oliveira; Lose, Alícia Duhá; Carneiro, Zenaide de Oliveira Novais; Cardoso, Lara da Silva. Testamento de André Vidal de Negreiros (1678).

¹⁴ Disponível em: <http://www.uefs.br/cedohs/view/coletaneas.html>.

¹⁵ A maior parte dos documentos trabalhados até agora está localizada em Portugal, no Arquivo Histórico Ultramarino (AHU), no Arquivo Nacional da Torre do Tombo, na Biblioteca Central da Universidade de Coimbra; mas também na Holanda, no Arquivo Nacional dos Países Baixos. Foi microfilmada e digitalizada, inicialmente, pelo Projeto Resgate Barão do Rio Branco. Em cada volume, consta a informação precisa de localização da documentação nele editada. Para os trabalhos da Série, todavia, análises físicas têm sido realizadas em contato direto com os documentos, e novas imagens têm sido produzidas, pelo que muito agradecemos aos arquivos mencionados, que têm compreendido a necessidade desse contato físico com os originais e das novas imagens, viabilizando um acesso, em nível de exceção.

Volume 3	Paschoal, Izaías Araújo das Neves; Carneiro, Zenaide de Oliveira Novais; Lose, Alícia Duhá; Lacerda, Mariana Fagundes de Oliveira. Documentos dos terços de homens pretos e pardos (1650-1793).
Volume 4	Brito, Rosana Carvalho; Carneiro, Zenaide de Oliveira Novais; Lose, Alícia Duhá; Lacerda, Mariana Fagundes de Oliveira. Documentos de Comissários baianos da Inquisição portuguesa no Brasil (1700-1791).
Volume 5	Santos, Elane Santos e; Lacerda, Mariana Fagundes de Oliveira; Lose, Alícia Duhá; Carneiro, Zenaide de Oliveira Novais. Documentos de Comissários baianos do Santo Ofício no Brasil (1698-1809).
Volume 6	Santos, Emília Carneiro dos; Lacerda, Mariana Fagundes de Oliveira; Lose, Alícia Duhá; Carneiro, Zenaide de Oliveira Novais; Cardoso, Lara da Silva. Documentos da família Vieira Ravasco (1617-1725).
Volume 7	Souza, Márcia Silva de; Carneiro, Zenaide de Oliveira Novais; Lose, Alícia Duhá; Lacerda, Mariana Fagundes de Oliveira; Cardoso, Lara da Silva. Documentos de mamelucos (1694-1723).
Volume 8	Lemos, Maiara da Silva; Carneiro, Zenaide de Oliveira Novais; Lose, Alícia Duhá; Lacerda, Mariana Fagundes de Oliveira. Documentos escritos por portugueses no Brasil (1551 e 1736).
Volume 9	Carneiro, Bruna Trindade Gomes; Lose, Alícia Duhá; Carneiro, Zenaide de Oliveira Novais; Lacerda, Mariana Fagundes de Oliveira. Um Códice em Língua Geral (1750-1758).
Volume 10	Alkimin, Ilma Magalhães; Lose, Alícia Duhá; Lacerda, Mariana Fagundes de Oliveira; Carneiro, Zenaide de Oliveira Novais. Sermões de Eusébio de Matos (1629-1692).

Volume 11	Souza, Ionáia Maria de Jesus; Lose, Alícia Duhá; Lacerda, Mariana Fagundes de Oliveira; Carneiro, Zenaide de Oliveira Novais. Cartas, requerimentos, pareceres e afins (1712-1758).
Volume 12	Jesus, Patrícia dos Santos; Lacerda, Mariana Fagundes de Oliveira; Lose, Alícia Duhá; Carneiro, Zenaide de Oliveira Novais; Cardoso, Lara da Silva. O Borrador de Antônio Gomes Ferrão Castelo Branco (1700-1800).
Volume 13	Silva, Alan Souza da; Lose, Alícia Duhá; Carneiro, Zenaide de Oliveira Novais; Lacerda, Mariana Fagundes de Oliveira. Documentos da Jornada do Rio Grande (1597-1598).
Volume 14	Dantas, Andreza Alves; Carneiro, Zenaide de Oliveira Novais; Lose, Alícia Duhá; Lacerda, Mariana Fagundes de Oliveira. Da documentação das Academias Brasileiras na Bahia Setecentista (1724-1725).
Volume 15	Carvalho, Gabriela Fortalesa de; Carneiro, Zenaide de Oliveira Novais; Lose, Alícia Duhá; Lacerda, Mariana Fagundes de Oliveira. Cartas selecionadas do padre jesuíta português, Manuel da Nóbrega (1517-1570), escritas no Brasil entre 1549-1560.
Volume 16	Soares, Daniela Oliveira de Souza; Carneiro, Zenaide de Oliveira Novais; Lose, Alícia Duhá; Lacerda, Mariana Fagundes de Oliveira. Correspondências entre José da Silva Lisboa e Balthasar da Silva Lisboa, escritas na Bahia, entre 1780 e 1805.

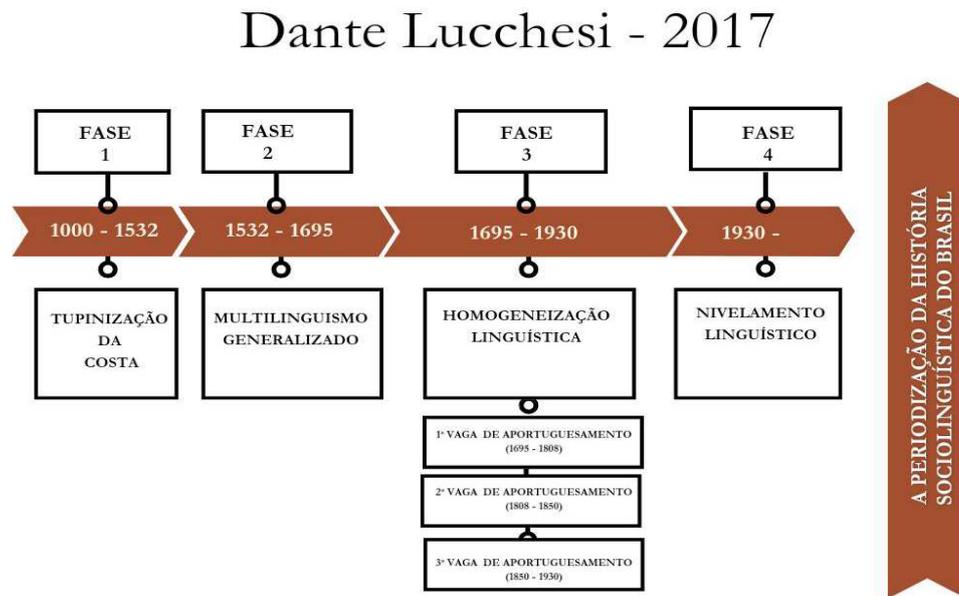
Fonte: Elaborado pelas autoras.

O trabalho de constituição desse *corpus* colonial do CE-DOHS, a nosso ver, representa, junto ao PHPB, um avanço de sua agenda de *corpus* diacrônico, tanto em razão da ampliação do banco de textos – procurando cobrir etapas da periodização da história sociolinguística do Brasil – como em razão do diálogo, com maior grau de maturidade, com a Diplomática, a Paleografia e a Filologia.

O CE-DOHS tem considerado a proposta de periodização da história sociolinguística do Brasil defendida por Lucchesi (2017)¹⁶ e sistematizada na Figura 4.

¹⁶ Conferir também Lucchesi e Callou (2020) sobre os cenários sociolinguísticos do Brasil colônia.

Figura 4: Proposta de periodização da história sociolinguística do Brasil, de Lucchesi (2017)



Fonte: Elaborado pelas autoras.

Conforme o autor,

[a]o lado de uma vasta pesquisa empírica, possibilitada pelo ingente trabalho de organização de *corpora* históricos, esse movimento também suscitou uma reflexão mais geral sobre a história da língua no Brasil, buscando compreender não apenas a formação da realidade linguística, mas também os fatores que lhe foram determinantes. É nessa perspectiva que se insere a proposta aqui formulada, pois uma periodização deve necessariamente relacionar-se a uma visão de conjunto da história sobre a qual se debruça. Quanto mais explícita for essa relação, mais a periodização alcançará seu objetivo maior: balizar a compreensão de conjunto da história que divide e orientar as pesquisas empíricas no campo. A periodização aqui apresentada busca, assim, contribuir para a compreensão da história da língua no Brasil, particularmente no que concerne aos seus condicionamentos sócio-históricos. Nesse sentido, adota-se aqui como ponto de partida a perfeita síntese de Rosa Virgínia Mattos e Silva (2004: 14), que definiu a história sociolinguística do Brasil como a passagem de “um multilinguismo generalizado a um multilinguismo localizado” (Lucchesi, 2017, p. 348-349).

Com a execução da fase 2 (colonial), o banco CE-DOHS vem cobrindo de forma mais satisfatória a fase 3 (da homogeneização linguística, conforme apontada por Lucchesi, 2017), com documentos dos séculos XVII e XVIII, que se somam aos documentos da fase 1 (pós-colonial) de execução do projeto Guarda-Chuva.

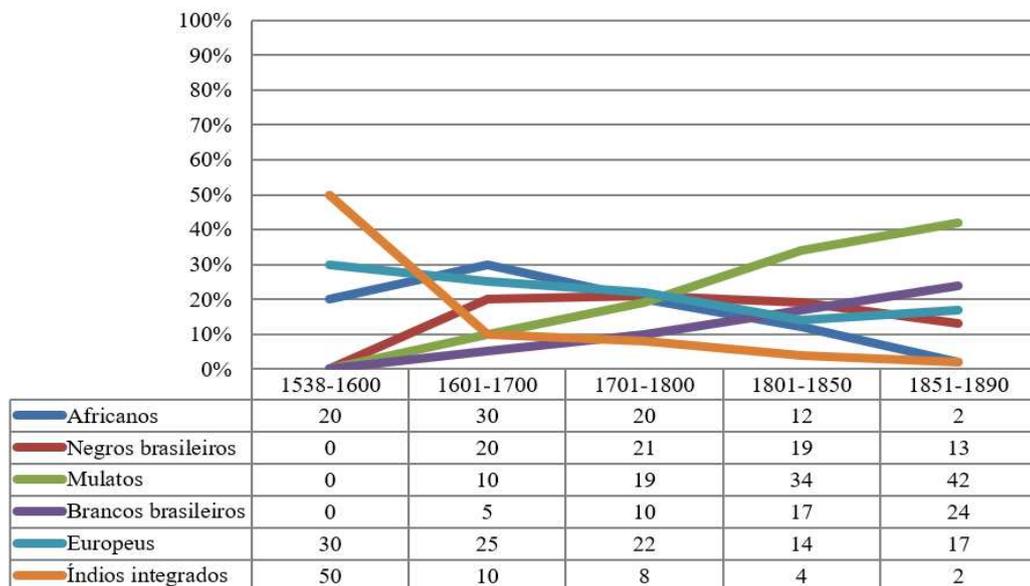
Considerando as proposições de Mattos e Silva – que participou da criação do PHPB e foi coordenadora do PHPB Bahia por muitos anos –, sintetizadas por Lobo

(2015, p. 70), o CE-DOHS/NELP/UEFS desenvolve suas pesquisas sobre a história social linguística do PB:

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

Além dessas proposições, o CE-DOHS/NELP/UEFS também considera dados da demografia histórica do Brasil (Mussa, 1991), que revelam os atores principais na cena linguística. O Gráfico 1, adaptado de Mussa (1991), apresenta as porcentagens de raça presentes nas diversas fases de colonização.

Gráfico 1 – Demografia histórica do Brasil. Adaptado de Mussa (1991, p. 163)¹⁷



Fonte: Lemos (2023, p. 29).

¹⁷ A respeito da tabela de Mussa, ressaltamos a importância de utilizar, nos dias atuais, diferentemente do autor, “miscigenado” em lugar de “mulato”, dado o teor racista presente neste termo, tanto em sua etimologia quanto em sua formação discursiva.

O NELP não temo *corpus* ideal nem trabalha com esse conceito. O que oferecemos é um banco, com lacunaridade¹⁸, para uma aproximação da história linguística dos diferentes grupos envolvidos nessa complexa formação do Brasil, sabendo que, no que diz respeito aos grupos sociais subalternizados, as lacunas são ainda maiores, haja vista que “estes raramente deixaram testemunhos autógrafos”, e a escrita de sua história linguística pauta-se, em muitos aspectos, “em uma reconstrução a partir de “indícios”, e, necessariamente, do presente em direção ao passado”¹⁹ (Lobo, 2001, p. 109). Como disse Mattos e Silva (2008a), trata-se de um trabalho de arqueologia da linguagem, pesquisa indiciária.

Na descrição do perfil sociocultural dos escreventes, entendendo a escrita a partir de quem a escreveu em um determinado contexto sócio-histórico, o CE-DOHS/NELP/UEFS vai além das categorias tradicionais de perfil, investindo esforços para o que Lopes *et al* (2010, p. 242) denominaram “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 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).

Esse cruzamento de olhares e cruzamento de perspectivas estão sempre sob o escopo teórico-metodológico da Paleografia e da Filologia, entendido como condição *sine qua non* para a formação de *corpora* em Linguística Histórica e no âmbito do PHPB.

Fazendo os controles sócio-histórico e paleográfico da documentação, o CE-DOHS descreve o “quem” e o “para quê” (Petrucci, 2003), além de descrever, no âmbito

¹⁸ Segundo Paixão de Souza (2006, p. 25-26), “a documentação possível sobre o passado das línguas chegou até nós através da escrita. Isso traz antes de tudo o problema da lacunaridade do registro escrito: de partida, o estudo das línguas sem tradição de escrita, naturalmente, não é possível por essa via; e tampouco é possível estudar, pela via do registro, todas as etapas da evolução dos idiomas com escrita, em vista das lacunas documentais”. A autora continua: “Ora, esses registros representam um fragmento dos acontecimentos. Mais que isso: um fragmento daquilo que um determinado contexto histórico julgou relevante registrar; que um segundo momento histórico julgou importante preservar; e que um terceiro momento histórico considerou pertinente examinar” (Paixão de Souza, 2006, p. 36).

¹⁹ No CE-DOHS, as coleções documentais representativas das normas socialmente estigmatizadas do PB são: Correspondências Amigas; Cartas em Sisal; Cartas Marienses, todas do século XX. Sobre as Cartas em Sisal, editadas por Santiago (2012), foi publicada uma coletânea de estudos morfofossintáticos, organizada por Santiago, Lacerda e Carneiro (2023), com financiamento do Programa de Auxílio Interno aos Programas de Pós-Graduação *Stricto Sensu* (AUXPPG/UEFS).

da Paleografia tradicional²⁰, o “quando” e o “onde”. São variáveis que, segundo Mattos e Silva (2004), amenizam, com o recurso também à **historiografia**, problemáticas metodológicas da Sociolinguística Histórica, como os problemas da Representatividade e da Autenticidade (Hernández-Campoy; Shilling, 2012).

É com grande esforço e uma equipe especializada em campos diversos que o PHPBBahia temse dedicado, ao longo de anos, à constituição de *corpus* diacrônico, entre outras agendas do PHPB nacional, sempre buscando aprimorar esse trabalho, dando aos *corpora* o tratamento teórico-metodológico necessário.

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; controles paleográfico e filológico; 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 Paleografia, a Filologia, a História Social da Cultura Escrita, a Linguística Computacional, entre outros (Santiago *et al*, 2021, p. 314).

Essas etapas precedem os estudos linguístico-gramaticais, pois “[é] nos textos remanescentes do passado, afinal, que encontramos os dados diacrônicos para a pesquisa em Linguística Histórica” (Mattos e Silva, 1998, p. 106).

5 O tratamento paleográfico-filológico no âmbito do CE-DOHS fase 2 (colonial): uma ilustração

A Paleografia é a ciência que estuda as escritas antigas, decifrando-as, lendo-as, interpretando-as e situando-as nas coordenadas espaço-temporais. A Paleografia Diplomática, por sua vez, é uma subdisciplina da Paleografia, que se concentra especificamente no estudo de documentos oficiais e administrativos, com o objetivo de analisar e interpretar a escrita e a formatação desses documentos, para entender melhor seu contexto histórico, jurídico e administrativo. Portanto, à Paleografia Diplomática interessam os elementos materiais usados para produção da escrita; os modelos caligráficos usados por diferentes comunidades de escrita e em determinados locais e períodos; as fórmulas textuais empregadas para variados tipos de texto, em contextos específicos; e tudo o mais que possa auxiliar na correta interpretação da escrita, na identificação de sua validade e veracidade,

²⁰ Conferir livro *Paleografia e suas interfaces*, organizado por Lose e Sacramento (2018).

na sua datação e na identificação das mãos que produzem os textos.

A Filologia, por sua vez, é a ciência que estuda as línguas e seus escritos, analisando e interpretando a evolução, estrutura e significado dos escritos ao longo do tempo. Uma parte importante da Filologia é a Ecdótica, ou Crítica Textual, que se concentra na preparação de edições de textos, a partir de critérios coerentes e adequadamente aplicados, fazendo com que um texto seja deslocado no tempo e no espaço, sem perder sua história.

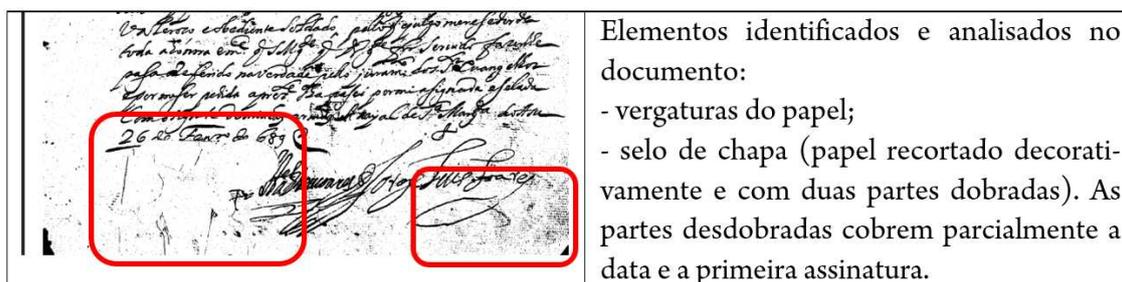
Para o trabalho que vem sendo realizado em parceria com a Linguística Histórica, temos aplicado, para seleção do material a ser transcrito e editado, as metodologias da Paleografia Diplomática, disciplina vital para historiadores da língua que trabalham com textos de épocas passadas, proporcionando ferramentas e métodos para a análise rigorosa e contextualização de registros antigos.

No caso da *série Documentos para a história linguística do Brasil Colônia*, já aqui referida, tomemos como exemplo o conjunto de textos do Terço dos Homens Pretos e Pardos, instituição encabeçada por Henrique Dias, na capitania de Pernambuco, no âmbito das contendas para expulsão dos holandeses das terras brasileiras. Nas mãos de Izaías Araújo das Neves Paschoal, sob a orientação conjunta de Zenaide de Oliveira Novais Carneiro e Alícia Duhá Lose,

os documentos manuscritos deixados pelos integrantes dos Terços de Homens Pretos e Pardos passaram por um extenso processo de edição semidiplomática e análise paleográfica – dos escritos e das mãos que levantaram a pena – de modo a garantir a originalidade e a confiabilidade para um estudo linguístico. Sem o apoio da Crítica Textual, da Diplomática e da Paleografia, não seria possível garantir a data de produção dos documentos, tampouco confirmarmos se o assinante fora, realmente, o *scriptor* responsável para que pudéssemos colher os dados linguísticos (Paschoal, 2024, p. 20).

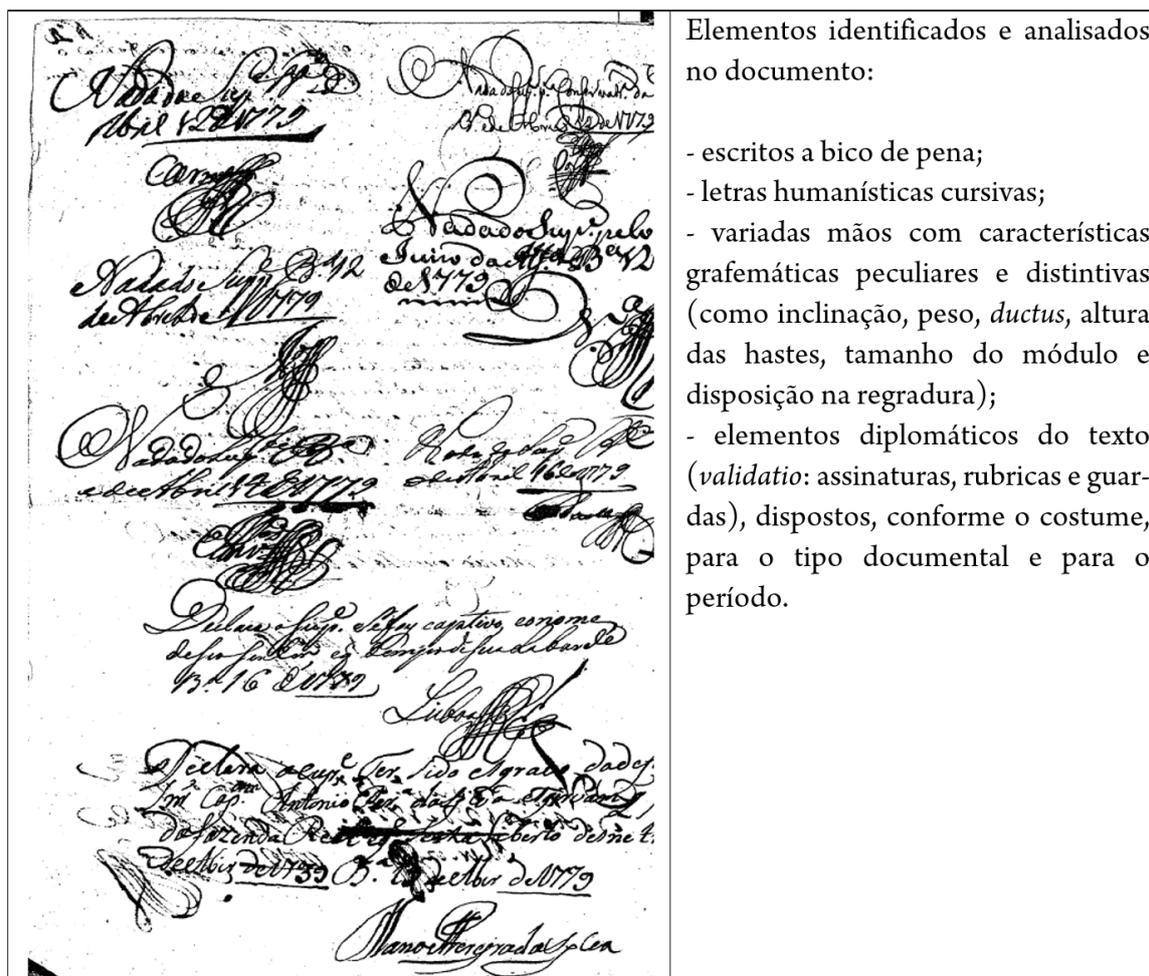
Todos os documentos trabalhados neste conjunto estão custodiados pelo Arquivo Ultramarino, em Lisboa, e foram disponibilizados através do Projeto Resgate Barão do Rio Branco. Todas as análises, por conseguinte, foram feitas a partir das imagens digitais dos microfilmes de cada documento. Não obstante, procedemos à **crítica** paleográfico-diplomática externa: aquela que foca nos aspectos materiais e físicos, analisando características do suporte, dos instrumentos de escrita, dos elementos constitutivos do traçado das letras, dos modelos caligráficos e da *mise-en-page*. Foi possível analisar também as estruturas diplomáticas dos textos e os elementos de validação neles presentes. As Figuras 5, 6 e 7, a seguir, apresentam a sistematização dessas informações.

Figura 5: Visualização das vergaturas do papel, provavelmente de trapoe do selo de chapa colado sobre a data e a assinatura



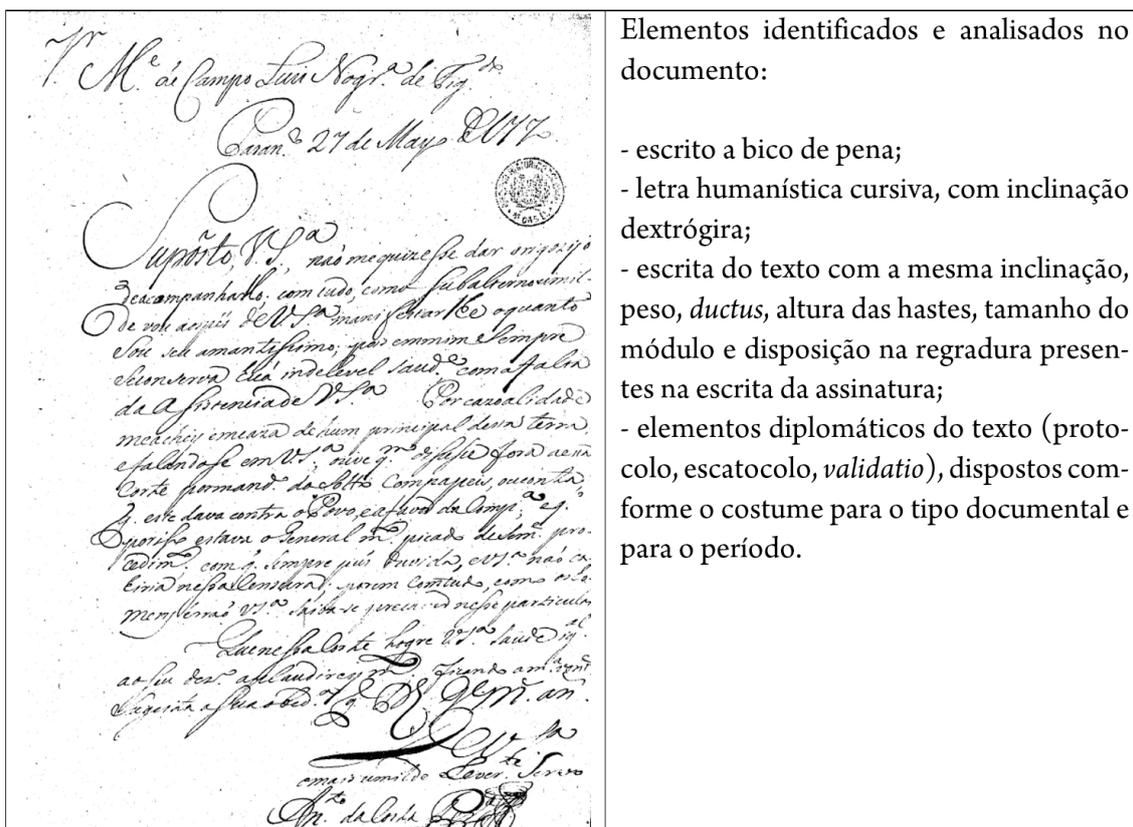
Fonte: ATHPP-JLS-RN-CER-1689.2

Figura 6: Assinaturas, rubricas e guardas, acompanhando os “nada consta” e as declarações



Fonte: ATHPP-ASJ-BA-ATE-1774

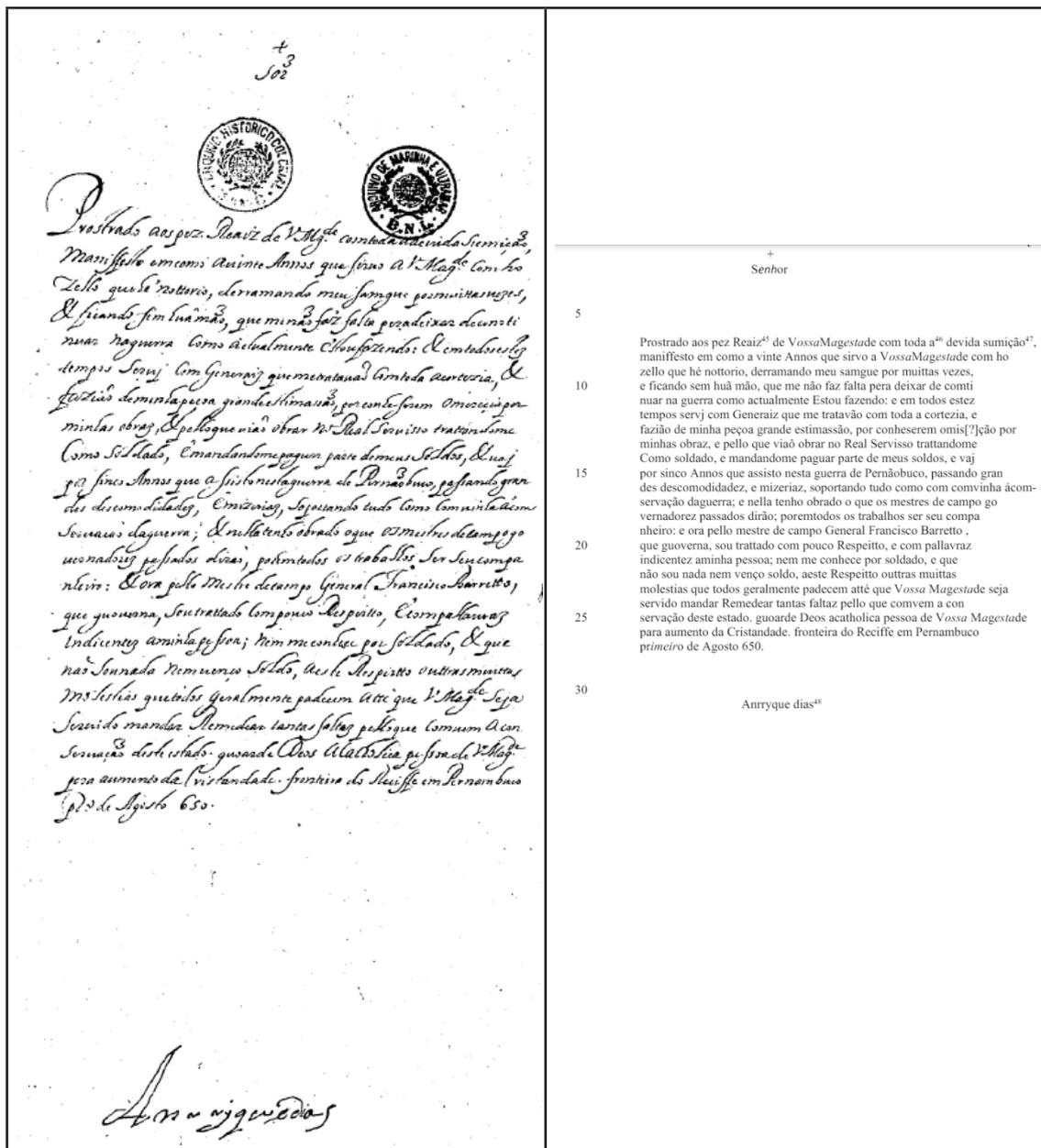
Figura 7: Elementos comprobatórios de autoria mecânica e intelectual



Fonte: ATHPP-ACP-PE-OFI-1770

Depois das análises paleográfico-diplomáticas, foram selecionados, com confiabilidade científica, 28 documentos notariais, manuscritos entre 1650 e 1793, no período do Brasil Colonial, por homens pretos e pardos que integravam os Terços de Homens Pretos e Pardos. Tais documentos foram lidos, transcritos e editados a partir de critérios conservadores: mantendo-se as características linguísticas do texto original, inclusive no procedimento de desenvolvimento de abreviaturas; as letras omitidas foram incluídas em destaque, obedecendo a forma estendida encontrada noutra parte do documento ou em documentos coetâneos.

Figura 8: Fac-símile e edição semidiplomática de carta datada de agosto de 1650, passada por Henrique Dias para Vossa Majestade, na qual descreve a má administração do Mestre de Campo Francisco Barreto, que o trata com palavras indecentes e não lhe paga o devido soldo.



Fonte: Elaborada pelas autoras, a partir de Paschoal (2024, p. 244).

A Figura 8 ilustra o trabalho de edição semidiplomática realizado por Paschoal (2024) na sua dissertação de mestrado, defendida e aprovada no PPGEL da UEFS, tendo uma banca examinadora composta pela professora Mariana Fagundes de Oliveira Lacerda e pelo professor e historiador Luiz Geraldo Silva.

6 Possibilidades de estudos linguísticos do *corpus* colonial do CE-DOHS

Uma vez garantida a autenticidade dos materiais – com esse trabalho conjunto da Paleografia, da Filologia e da Linguística Histórica –, que é uma exigência decisiva (Maia,

2012, p. 538), eles podem ser explorados, em diferentes níveis de análise e diferentes quadros teóricos, seja na perspectiva da Linguística Histórica Sócio-Histórica, seja na perspectiva da Linguística Diacrônica (Mattos e Silva, 2008a).

Em relação ao *corpus* colonial do CE-DOHS, Cardoso, Carneiro e Lacerda (2021, p. 350-351) listaram as seguintes possibilidades de estudos linguísticos, entre outras:

Estudo de painel (Labov, 1994; Rumeu, 2008); estudo de tendências (Labov, 1994); análises contrastivas entre acervos; contraste com *corpora* portugueses do mesmo período; contraste entre as normas usadas pelos colonos e as normas linguísticas propostas pelos portugueses para uso da LP no século XVIII; estudo das marcas de inabilidade dos manuscritos do *corpus*; estudo de fenômenos linguísticos diversos, visando à identificação de marcas específicas do português do Brasil Colonial; ampliação do *corpus* diferencial do PHPB (Cardoso; Carneiro; Lacerda, 2021, p. 350-351).

O CE-DOHS fase 2 (colonial), com seis anos de execução, já apresenta alguns frutos, entre eles: dissertações, teses, artigos, capítulos, a coleção *O Sertão por escrito* e, em breve, a série de documentos e coletâneas de estudos linguístico-gramaticais. Mas nada disso seria possível sem aqueles que vêm discutindo, com espírito científico e ética, a questão do *corpus* no PHPB desde muito tempo. Nesse contexto, destacamos o papel de Ilza Ribeiro desde a criação do NELP/UEFS, com seu projeto inaugural *Contribuições para a Constituição de um Banco de Textos e de um Banco de Dados para o Estudo da História do Português do Brasil, do Século XVII ao XX*.

A respeito das pesquisas de mestrado e doutorado, já foram defendidas cinco dissertações e cinco teses; em andamento, acham-se duas dissertações e duas teses. É importante destacar que todos esses estudos (14, no total)²¹ atendem a três agendas do PHPB (constituição de *corpus* diacrônico, estudo sócio-histórico e estudo linguístico); e seguem um mesmo modelo de elaboração: os mesmos tratamentos diplomático, paleográfico, filológico e sócio-histórico. O estudo linguístico realizado segue uma orientação de natureza sobretudo qualitativa, feito em uma abordagem descritivo-interpretativa, comum em Linguística Histórica e que se trata de uma sistematização metodológica que busca “apresentar uma descrição organizada dos factos lingüísticos”, tendo em vista “uma gramática descritiva, que opere sobre inventários que se definam como representativos” (Mattos e Silva, 1989, p. 44).

²¹ Encontra-se em organização a série *Dissertações e teses para a história sociolinguística do Brasil*, com a publicação de pesquisas que atendem o projeto Guarda-Chuva CE-DOHS: um *corpus* para uma caracterização linguístico-gramatical do português brasileiro - fase colonial e fase pós-colonial. Nessa série, serão apresentados os resultados de exploração dos *corpora* editados, tanto a exploração paleográfico-filológica como as explorações sócio-histórica e linguístico-gramatical.

Entre os dez trabalhos já defendidos, seis abordam fenômenos gramaticais: as dissertações de Santos (2022), Araújo (2023) e Paschoal (2024) e as teses de Brito (2024), Santos (2024) e Lemos (2025)²²; as dissertações de Santos (2019) e Souza (2022) e as teses de Silva (2021) e Carneiro (2025) abordam índices grafo-fonéticos²³.

Sobre os seis estudos descritivo-gramaticais já defendidos no PPGEL da UEFS e como parte do CE-DOHS fase 2 (colonial), foram selecionados os seguintes fenômenos: a concordância verbal, a sintaxe das orações relativas, o complemento verbal *lhe*, investigados nas dissertações de Santos (2022), Araújo (2023) e Paschoal (2024), respectivamente; e os sintagmas possessivos, os dativos e os clíticos, investigados nas teses de Brito (2024), Santos (2024) e Lemos (2025), respectivamente. Durante o desenvolvimento de suas pesquisas, os seis autores enfrentaram o problema da invariação, em maior ou menor grau, apontados por Hernández-Campoy e Schilling (2012) como um dos principais problemas teórico-metodológicos em estudos linguísticos sócio-históricos. Esse problema acontece porque, em geral, textos escritos fornecem pouca possibilidade de variação linguística, sobretudo quando o *corpus* possível são documentos notariais, como é o caso dos documentos explorados nas três dissertações em questão.

A pesquisa de Araújo (2023), por exemplo – orientada por Mariana Fagundes de Oliveira Lacerda e Alicia Duhá Lose –, cujo foco é a sintaxe das relativas no Testamento de André Vidal de Negreiros, do século XVII, (tema que se destaca, sendo considerado decisivo na emergência do português brasileiro (Tarallo, 1983, 1993), em um processo de “garimpar”, chegou aos seguintes resultados, localizando três ocorrências de relativas não-padrão. São eles:

Embora tenham sido levantadas, no total, 153 sentenças relativas, a maior parte desses números se trata do emprego de estratégias de relativização padrão. Grande parcela dessa constatação deve-se à quantidade de registros de cláusulas cujas funções relativizadas são as de Sujeito e de Objeto Direto, que, juntas, alcançam um total de 110 ocorrências padrão. Além disso, nas posições mais baixas, também são predominantes tais estratégias. (...) São apenas três as sentenças relativas de estratégia não-

²² Resultados desses estudos gramaticais, entre outros, fazem parte do livro, intitulado *Oito estudos gramaticais sobre o período colonial brasileiro* (no prelo, Pontes Editores), sob a organização de Mariana Fagundes de Oliveira Lacerda, Zenaide de Oliveira Novais Carneiro e Elane Santos e Santos, com recursos SEPESQ/UEFS, e do livro, intitulado *Para a história do português no Brasil colonial: revisitando problemas e ampliando análises* (no prelo, UEFS Editora), sob a organização de Rosana de Carvalho Brito, Elane Santos e Santos, Mariana Fagundes de Oliveira Lacerda e Zenaide de Oliveira Novais Carneiro, com recursos SEPESQ/UEFS.

²³ Resultados dessas pesquisas de natureza grafo-fonética fazem parte do livro, intitulado *Corpora histórico-diacrônicos do português brasileiro: da descrição gramática aos indícios da oralidade* (no prelo, Pontes Editores), sob a organização de Huda da Silva Santiago, Adilson Silva de Jesus, Mariana Fagundes de Oliveira Lacerda e Zenaide de Oliveira Novais Carneiro, com recursos SEPESQ/UEFS.

padrão identificadas no *corpus*. Duas delas cuja função relativizada é a de Objeto Direto e apenas uma cuja função relativizada é a de Adjunto (Araújo, 2023, p. 105-106).

Araújo (2023) evidencia que “vestígios de tinta” de relativas não-padrão, como as relativas cortadoras e as relativas resumptivas, podem ser encontrados em textos escritos durante o Brasil Colonial. É mesmo um trabalho de “formiga obreira”, aqui e ali atrás de dados, em busca de indícios que ajudem a contar a história do PB. Os *corpora* de Santos (2022), Araújo (2023) e Paschoal (2024) são datados dos séculos XVII e XVIII, quando ainda não era possível atestar a existência do português do Brasil; mas investigar como se configurava o português no Brasil pode contribuir para uma reconstrução, por aproximação, da história social linguística do PB. É nisto que o CE-DOHS tem investido esforços desde 2017: enfrentar a raridade documental, quando se trata do período colonial no Brasil, e procurando fazer o melhor uso dos dados disponíveis.

7 Considerações finais: o trabalho com *corpus* não é cômodo, mas é necessário

Desde o início do projeto, em 1997, Mattos e Silva adverte a todos sobre a “enormidade do empreendimento” (Castilho, 2004, p. 3), referindo-se ao PHPB. Além disso, sobre a agenda de *corpus* diacrônico do projeto, a autora complementa:

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

A fala de Mattos e Silva (2004) destaca a importância de fazer um trabalho que não é cômodo, mas que é necessário, como necessário é o diálogo entre a Linguística Histórica, a Diplomática, a Paleografia e a Filologia. Sem esse diálogo não é possível o trabalho com *corpus* para estudos linguísticos e discussões teóricas com solidez.

Ao longo de 28 anos, os pesquisadores do PHPB, nas diferentes fases de trabalho, “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), têm buscado executar as agendas de trabalho desse enorme empreendimento, trazendo, com cada equipe regional, como a equipe da Bahia, elementos *Para uma História do Português Brasileiro*. Elementos esses que sempre podem, numa perspectiva interdisciplinar, como defendemos aqui, ser revisitados e refinados.

Como “quixotes ou como loucos” (Mattos e Silva, 2004, p. 7) ou “como formigas obreiras” (Mattos e Silva, 2008b, p. 40), os pesquisadores do PHPB continuam empenhando esforços no seu grande empreendimento... E o NELP/UEFS tem sido um braço forte, especialmente, mas não apenas, na agenda de constituição de *corpus* diacrônico; como afirma Castilho, no prefácio da série colonial (no prelo), “O NELP, com quase 30 anos de trabalho intenso, tem já uma enorme folha de serviços em favor da Linguística Histórica no Brasil.”

Referências

ARAÚJO, V. S. S. *O Testamento de Vidal de Negreiros (século XVII): edição semidiplomática e elementos para uma descrição da sintaxe das relativas*. Dissertação (Mestrado em Estudos Linguísticos). Feira de Santana, UEFS, 2023.

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 abril 2004.

BRITO, R. C. *Pelas mãos de Comissários do Santo Ofício na Bahia setecentista: edição semidiplomática e estudo do artigo definido diante de pronome possessivo*. Tese (Doutorado em Estudos Linguísticos). Feira de Santana, UEFS. 2024.

CARDOSO, L. S.; LACERDA, M. F. O.; CARNEIRO, Z. O. N. Relações entre a constituição de corpora diacrônicos do português brasileiro e os princípios da Ciência Aberta. *ABRALIN em Cena 17: dados linguísticos*, 2023.

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.

CARNEIRO, Z. O. N.; LACERDA, M. F. O. (Org). *CE-DOHS - Corpus Eletrônico de Documentos Históricos do Sertão - Início*. Disponível em: <http://www.uefs.br/cedohs/view/home.html>. Acesso em: 14 março 2024.

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.; SILVA, A. J.; SOUZA, E. H. P. M. *Livro de Razão (XVIII-XIX): edições fac-similar e semidiplomática*. Feira de Santana: Editora UEFS, 2022, v.1.

CARNEIRO, Z. O. N.; LACERDA, M. F. O.; LOSE, A. D. (Coordenação Geral). *Série Documentos para a HistóriaLinguística do Brasil Colônia*. Feira de Santana: UEFS Editora (No prelo).

CARNEIRO, B. T. G. *Um códice em língua geral: edição, estudo paleográfico e sócio-história da Amazônia (1750-1758)*. Tese (Doutorado em Estudos Linguísticos). Feira de Santana, UEFS. 2025.

CASTILHO, A. T. Historiando a língua portuguesa no novo mundo. In: *XX Conferência final no Congresso Internacional da Associação Brasileira de Linguística e Filologia da América Latina (ALFAL), Projeto 3 - História do Português Brasileiro, da Europa à América*. Faculdade de Humanidades e Arte. Universidade de Concepción, Chile, 2024.

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: 1 março 2023.

DUARTE, M. E. L.&CALLOU, D. M. I. (Org.). *Para a História do Português Brasileiro*, v. IV: Notícias de corpora e outros estudos. Rio de Janeiro: Universidade Federal do Rio de Janeiro / Faperj, 2002.

HERNÁNDEZ-CAMPOY, Juan M.; SCHILLING, Natalie. The Application of the Quantitative Paradigm to Historical Sociolinguistics: problems with the generalizability principle. In: HERNÁNDEZ-CAMPOY, Juan Manuel; CONDE-SILVESTRE, Juan Camilo. *The Handbook of Historical Sociolinguistics*. New Jersey: Blackwell Publishing, 2012. p. 61-79.

LACERDA, M. F. O. O Núcleo de Estudos de língua Portuguesa da Universidade Estadual de Feira de Santana: centro de Linguística Histórica no Semiárido baiano. In: *Revista Confluência*. 2024. p. 147-199.

LACERDA, M. F. O.; CARNEIRO, Z. O. N.; LOSE, A. D. O trabalho de organização de corpora históricos, a pesquisa empírica e a periodização da história sociolinguística do Brasil. In: *30ª Jornada do GELNE*. Universidade Estadual de Feira de Santana. Feira de Santana, 2024.

LACERDA, M. F. O.; CARNEIRO, Z. O. N.; LOSE, A. D. O CE-DOHS na rede de pesquisa PHPB: recuando ao período colonial. In: *XI Seminário de Estudos Filológicos: Filologia e Ensino: perspectivas críticas e editoriais*. Universidade Estadual de Feira de Santana, Feira de Santana, 2024.

LACERDA, M. F. O.; CARNEIRO, Z. O. N. (Org.). *Núcleo de Estudos de Língua Portuguesa*.

Disponível em: <https://nelp.uefs.br/>. Acesso em: 14 mar 2024.

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.

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 (Doutorado em Estudos Linguísticos). Feira de Santana, UEFS. 2023.

LEMOS, M. da S. *A língua vinda d'álém-mar: colocação de clíticos na escrita de portugueses no Brasil colonial*. Tese (Doutorado em Estudos Linguísticos). Feira de Santana, UEFS. 2025.

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. L. Arquivos, acervos e a reconstrução histórica do português brasileiro. In: OLIVEIRA, K. CUNHA E SOUZA, H.; SOLEDADE, J. *Do português arcaico ao português brasileiro: outras histórias*. Salvador, Edufba, 2009. p. 305-327.

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. *Grupo de Pesquisa Modus Scribendi- Grupo de Pesquisas Paleográficas, Filológicas e Históricas*. Disponível em: <https://letras.ufba.br/modus-scribendi-grupo-de-pesquisas-paleograficas-filologicas-e-historicas>. Acesso em: 1 mar 2024.

LOSE, A. D. *Centro de Estudos, Pesquisa e Documentação Paleográfica do Memória e Arte (CEPEDOP)*. Disponível em: <https://www.memoriaarte.com.br/>. Acesso em: 1 mar2024.

LOSE, A. D.; LACERDA, M. F. O.; CARNEIRO, Z. O. N.; ARAÚJO, V. S. S. Fontes linguísticas para estudos do português no Brasil colonial: o exemplo do Testamento de André Vidal de Negreiros (séc. XVII). In: *IX Simpósio Mundial de Língua Portuguesa*. Universidade da Madeira, Ilha da Madeira. Portugal, 2024.

LOSE, A. D.; CARNEIRO, Z. O. N.; LACERDA, M. F. O.; CARDOSO, L. S. O português no Brasil colonial: reflexões sobre as edições de manuscritos históricos brasileiros. *XX Conferência final no Congresso Internacional da Associação Brasileira de Linguística e Filologia da América Latina (ALFAL), Projeto 3 - História do Português Brasileiro, da Europa à América*. Faculdade de Humanidades e Arte. Universidade de Concepción, Chile, 2024.

LOSE, A. D.; LACERDA, M. F. O.; CARNEIRO, Z. O. N. *A crítica paleográfica como base da seleção dos corpora*. In: *XI Seminário de Estudos Filológicos: Filologia e Ensino: perspectivas críticas e editoriais*. Universidade Estadual de Feira de Santana, Feira de Santana, 2024.

LOSE, A. D.; LACERDA, M. F.; CARNEIRO, Z. O. N. Reflexões sobre as edições de manuscritos históricos brasileiros: a relação entre o olhar paleográfico e o olhar linguístico. In: *Simpósio da Contemporaneidade à Antiguidade: temas em debate*. Centro de Estudos Clássicos e Humanísticos da Universidade de Coimbra, Coimbra, 2024.

LOSE, A. D.; LACERDA, M. F.; CARNEIRO, Z. O. N. Edições de manuscritos históricos brasileiros: a série Documentos para a história linguística do Brasil colônia. In: *VI Congresso Internacional de Linguística Histórica*. Faculdade de Letras da Universidade de Coimbra, Coimbra, 2023.

LOSE, A. D.; SOUZA, SACRAMENTO, A. de. Para uma filologia na pesquisa em linguística histórica. *Letras. Revista do Programa de Pós-Graduação em Letras da Universidade de Santa Maria, Santa Maria*, n. 60, 2022. p. 11-31. Disponível em: <https://doi.org/10.5902/2176148542058>. Acesso em: 3 abril 2023.

LOSE, A. D. L.; SACRAMENTO, A. *Paleografia e suas interfaces*. Salvador: Memória e Arte, 2018.

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

LUCCHESI, D.; CALLOU, D. Os cenários sociolinguísticos do Brasil colonial. In: Castilho, A. (Coordenador Geral). *História do português brasileiro*. v. 9. CALLOU, D.; LOBO, T. *História social do português brasileiro: da história social à história linguística*. São Paulo: Contexto, 2020. p. 156-181.

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 2024.

MAIA, C. Linguística Histórica e Filologia. In: LOBO, T.; CARNEIRO, Z.; SOLEDADE, J. et al. (Org.). *Rosae linguística histórica, história das línguas e outras histórias*. Salvador: EDUFBA, 2012. p. 533-541.

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.

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

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

MATTOS E SILVA, R. V. Teorias da mudança linguística e a sua relação com a(s) história(s) da(s) língua(s). In: *Revista de Estudos Lingüísticos da Universidade do Porto*. v. 3. 2008b. Disponível em: <https://ler.letras.up.pt/uploads/ficheiros/6874.pdf>. Acesso em: 1 abril 2024.

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

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

OLIVEIRA, K. (2006). *Negros e escrita no Brasil do século XIX: sócio-história, edição filológica de documentos e estudo linguístico*. Tese de Doutorado. Programa de Pós-Graduação em Letras e Linguística. UFBA, 2006.

PASCHOAL, I. A. N. *Documentos dos Terços de Homens Pretos e Pardos (1650-1793): edição semidiplomática e estudo do complemento verbal*. Dissertação (Mestrado em Estudos Linguísticos). Feira de Santana, UEFS, 2024.

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

RIBEIRO, I. M. *Constituição de um Banco de Textos e de um Banco de Dados para o Estudo da História do Português do Brasil, do Século XVII ao XX*. 1997. Disponível em: <https://nelp.uefs.br/o-nelp/historico/>. Acesso em: 13 abril 2024.

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

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. C. *Documentos da Família Vieira Ravasco (XVII-XVIII): edição semidiplomática e elementos para uma descrição da concordância verbal*. Dissertação (Mestrado em Estudos Linguísticos). Feira de Santana, UEFS, 2022.

SANTOS, E. S. *Entre Linhas e Testemunhos*: edição semidiplomática de manuscritos de comissários baianos do Santo Ofício e estudo descritivo de argumentos dativos. Tese (Doutorado em Estudos Linguísticos). Feira de Santana, UEFS. 2024.

SANTOS FILHO, L. *Uma comunidade rural do Brasil antigo*: aspectos da vida patriarcal no sertão da Bahia nos séculos XVIII e XIX. São Paulo: Editora Nacional, 1956.

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.

SOUZA, M. S. de. *Documentos de Mamelucos no Brasil Colonial*: edição e descrição dos aspectos gráficos. Dissertação (Mestrado em Estudos Linguísticos). Feira de Santana (UEFS). 2022.

TARALLO, Fernando. Diagnosticando uma gramática brasileira: o português d'aquém e d'além-mar ao final do século XIX. In: ROBERTS, I.; KATO, M. (Orgs.). *Português brasileiro: uma viagem diacrônica*. Campinas: Editora da UNICAMP, 2018 [1993].

TARALLO, Fernando. *Relativization strategies in Brazilian Portuguese*. Tese (Doutorado em Linguística) – University of Pennsylvania, Philadelphia, 1983.

