

NAS VEREDAS DO LÉXICO HISTÓRICO DO PORTUGUÊS BRASILEIRO

ON THE PATHS OF THE HISTORICAL LEXICO OF BRAZILIAN PORTUGUESE

Vanderci de Andrade Aguilera¹

Celciane Alves Vasconcelos²

Fabiane Cristina Altino³

RESUMO

Este artigo tem como propósito apresentar o percurso do Léxico Histórico do Português Brasileiro (LHisPB) desde a sua gênese até a sua concretização. Trata-se de pesquisa desenvolvida no âmbito do Projeto *Para a História do Português Brasileiro* (PHPB) e, ao contrário das demais obras que compõem a série *História do Português Brasileiro*, publicada sob a coordenação do Dr. Ataliba T. de Castilho no formato impresso, o LHisPB vem à luz no formato eletrônico e disponível para o público em geral na plataforma da Universidade Estadual de Londrina. O LHisPB compõe-se, atualmente, dos seguintes Léxicos Estaduais: Paraná, São Paulo, Rio de Janeiro, Bahia, Paraíba, Pernambuco e Rio Grande do Sul. Inspirado na obra de Company y Mellis (2002), *Léxico Histórico del Español de México*, os verbetes de cada uma das unidades estaduais foram construídos pelos itens lexicais de *corpora* históricos datados dos séculos XVII a XIX coletados pelas equipes estaduais do PHPB.

PALAVRAS-CHAVE: Léxico Histórico. Português Brasileiro. Metodologia. Resultados.

ABSTRACT

The purpose of this article is to present the trajectory of the Historical Lexicon of Brazilian Portuguese (LHisPB) from its genesis to its implementation. This is research developed within the scope of the Project for the History of Brazilian Portuguese (PHPB) and, unlike the other works that make up the History of Brazilian Portuguese series, published under the coordination of Dr. Ataliba T. de Castilho in printed format, the LHisPB comes to light in electronic format and available to the general public on the platform of the State University of Londrina. The LHisPB is currently composed of the following State Lexicons: Paraná, São Paulo, Rio de Janeiro, Bahia, Paraíba, Pernambuco and Rio Grande do Sul. Inspired by the work by Company y Mellis (2002), *Lexicon Histórico del Español de México*, the entries of each of the state units were constructed by lexical items from historical corpora dating from the 17th to the 19th centuries collected by the PHPB state teams.

KEYWORDS: Historical Lexicon. Brazilian Portuguese. Methodology. Results.

Introdução

A construção do Léxico Histórico do Português Brasileiro - LHisPB, coordenado por Vanderci Aguilera e Fabiane Altino, nasceu como sugestão do Dr. Ataliba de Castilho, idealizador do Projeto

¹ Universidade Estadual de Londrina (UEL), vanderci@uel.br, <https://orcid.org/0000-0003-3052-3710>.

² Universidade Estadual de Londrina (UEL), celciane@uel.br, <https://orcid.org/0000-0002-1600-1433>.

³ Universidade Estadual de Londrina (UEL), fabiane_altino@uol.com.br, <https://orcid.org/0000-0002-5987-2028>.

Para a História do Português Brasileiro, à equipe do Paraná. O LHisPB corresponde ao 12º volume da obra História do Português Brasileiro.

Com sede na Universidade Estadual de Londrina, o projeto teve início no ano de 2009 com a proposta de compor um banco de dados lexicográficos referentes aos períodos Colonial e Imperial do Brasil para fins de estudos diacrônicos e sincrônicos do léxico do Português Brasileiro, integrando, assim, a coletânea organizada e publicada sob a coordenação geral de Castilho (2019, 2020). O LHisPB vem na esteira de outras obras de cunho lexicográfico para complementar trabalhos já publicados, como os de Cunha (1978, 2006), Silva (2008) e Murakawa (2021).

O *Dicionário histórico das palavras portuguesas de origem tupi*, de 1978 e o *Vocabulário Histórico-Cronológico do Português Medieval* (2006), ambos de Antonio Geraldo Cunha, referências basilares para a composição de um Léxico, configuram-se em um levantamento exaustivo do léxico da língua portuguesa nos séculos XIII, XIV e XV (ALTINO; AGUILERA, 2010).

O *Dicionário da língua portuguesa medieval* (SILVA, 2008) contempla 17000 verbetes coletados junto a dezenas de fontes medievais já transcritas, como as *Cantigas de Santa Maria*, as *Décadas*, O *Cancioneiro da Ajuda*, As *Cantigas D'El rei Dom Dinis*, entre outras. Essas obras contribuem para o conhecimento do português falado no Brasil e abrem caminho para a contribuição de outros projetos que venham somar forças para a constituição da história do português brasileiro.

É esse o desejo do LHisPB: juntar-se a outros estudos dessa natureza, como o *Dicionário histórico do português do Brasil séculos XVI, XVII e XVIII*, coordenado por Murakawa e recentemente publicado, com a base de textos do português colonial, constituindo o *corpus* de referência para a elaboração do dicionário histórico do português colonial e os diversos trabalhos no âmbito da pós-graduação *stricto sensu*, concluídos ou em andamento, que buscam contribuir para o melhor conhecimento da língua portuguesa em território brasileiro. Soma-se a esses trabalhos a contribuição que certamente prestará o LHisPB para o estudo do Português do Brasil – PB.

1. História do LHisPB: Uma proposta para o volume XII

O Léxico Histórico do Português Brasileiro – LHisPB, cujo principal objetivo é apresentar, sob tratamento lexicográfico, todo o conteúdo lexical dos documentos manuscritos datados dos séculos XVII a XIX, que foram coletados, transcritos e editados pelas equipes, necessitou de tempo de maturação para que delineássemos o perfil do seu suporte – *on-line* e de alimentação de dados, o recorte no tempo e as normativas necessárias a esse labor.

Para a composição do LHisPB, como qualquer trabalho desse porte, o recorte do *corpus* foi inevitável devido à impossibilidade de abarcar tudo o que foi escrito entre os séculos XVII e XIX. Para tanto, seguimos o exemplo da obra de Company e Melis (2002) que, para o Léxico Histórico del Español de México, trabalharam com 320 documentos do período colonial mexicano, procedentes de duas fontes documentais: o Archivo General de Indias en Sevilla e o Archivo General de la Nación en la Ciudad de México.

Tendo essa obra por referência e analisando os documentos disponíveis para o LHisPB, a equipe do Paraná, responsável pela elaboração e manutenção do *site*, optou pela composição de *corpus* constituído por documentos diversos. É certo que Company e Melis (2002) afirmam terem dado preferência, quase absoluta, a documentos informais, cartas pessoais, denúncias e testemunhos em juízo, petições e informes de particulares, com o fim de buscar uma aproximação maior da língua falada do período colonial hispânico. A decisão de trabalhar apenas com cartas pessoais e documentos não oficiais no LHPB nos distanciaria, sobremaneira, da estrutura e concepção da obra na versão brasileira e reduziria, consideravelmente, o *corpus*.

Ao contrário do *Léxico histórico del español de México*, no LHisPB, não foi possível fazer a distinção entre texto de natureza oficial e texto de caráter coloquial (familiar, pessoal). Essa decisão se fundamenta na observação do *corpus* constituído para o projeto do Léxico Histórico do Paraná, pois, pela leitura mais acurada dos manuscritos paranaenses, concluímos que, apesar de oficiais (documentos notariais), a ortografia e a gramática ali aplicadas estão longe de ser consideradas a padrão. Esse fato foi atestado por Almeida-Baronas (2006, 2009), Toniolo (2008), Souza (2012), Assis (2008, 2013) e Bençal (2018) em seus estudos, confirmando que, pelo contrário, a forma expressa em muitos dos manuscritos está muito próxima de uma modalidade de língua provavelmente falada por boa parte dos brasileiros no tempo da Colônia e do Império.

A análise de correspondências oficiais dos séculos XVII, XVIII e XIX, emanadas das antigas vilas paranaenses, tem, pois, indicado que, na maioria dos casos, os escribas não parecem ter tanta habilidade gramatical e gráfica como se espera de documentos oficiais. Essa afirmação pode ser exemplificada com inúmeros casos de metáteses e hipéteses; oscilação ortográfica no mesmo documento e pelas mesmas mãos; liberdade gramatical na concordância verbal e nominal e na colocação pronominal.

Por outro lado, diante da diversidade de gêneros coletados pelas equipes, do grande volume de material já disponível e, tentando compor um *corpus* geral menos heterogêneo, optamos por selecionar dois tipos de documentos manuscritos: os oficiais (correspondências eclesiásticas, notariais, cartas ao governo, por exemplo); e os de caráter coloquial (familiar, pessoal, petições, entre outros). É importante observar, nesse particular, que nem todas as regionais editaram documentos dos mesmos gêneros de texto. Por exemplo: o Paraná trabalha atualmente com cerca de 1000 documentos oficiais e apenas uma dezena de cartas pessoais. Não cogitamos também, por ora, incluir peças de teatro, que foram objeto de estudos de sintaxe para a equipe do Rio de Janeiro. Assim sendo, foi necessário excluir: notícias de jornais, processos, peças de teatro, embora o indiscutível reconhecimento da importância de tais documentos para a história do PB.

Outro ponto crucial para o trabalho com os documentos foi o estabelecimento de um recorte temporal. Quase todas as regionais segmentaram os documentos com base no início e término de cada século. O estabelecimento de uma periodização justifica-se, pois é no período colonial (começo do século XVII) que se inicia a expansão do território brasileiro para além dos limites do Tratado de

Tordesilhas, em que bandeirantes paulistas saem do litoral, dirigem-se para o interior e atacam as reduções (também conhecidas como aldeamentos ou missões) estabelecidas pelos jesuítas espanhóis. Ainda no século XVII, entre invasões (holandesa e francesa) e a expansão canavieira no Nordeste, o Brasil é dividido em duas Colônias distintas: Estado do Brasil e Estado do Maranhão (posteriormente, Estado do Grão-Pará e Maranhão).

Durante o século XVIII, o cenário é o das guerras e revoltas por conquistas de território, a descoberta de diamantes em Minas Gerais, o Tratado de Madri anulando o Tratado de Tordesilhas, a expulsão dos jesuítas de Portugal e colônias, a laicização do ensino no Brasil, a extinção das últimas capitanias hereditárias, transformadas em capitanias reais, a transferência da capital do Brasil – de Salvador para o Rio de Janeiro - a reunificação dos Estados do Brasil e do Grão-Pará e Maranhão, a morte de D. José I e a ascensão de D. Maria I ao trono português, a consequente proibição no Brasil do funcionamento de quaisquer indústrias ou manufaturas (excetuando a água-ardente e tecidos grosseiros) e a Inconfidência Mineira. Esses são alguns dos acontecimentos que marcaram o centenário setecentista.

Por fim, o século XIX é pontuado por acontecimentos importantes para a constituição da nação brasileira. Entre eles destacamos a chegada de D. João VI e a corte portuguesa ao Brasil, dando início ao período joanino (1808/21); a abertura dos portos brasileiros, o alvará de Liberdade Industrial, revogando as proibições impostas por D. Maria I; a instalação da Imprensa Régia e publicação do primeiro jornal brasileiro; a criação de escolas de Medicina em Salvador e Rio de Janeiro – primeiros cursos superiores; a abolição da escravidão; a vinda de fortes contingentes imigrantes europeus e, encerrando esse século de guerras e revoluções, a proclamação da República. Uma vez justificada a segmentação da fonte documental, feita a partir do início e término de cada século, o LHisPB previu o tratamento dado aos documentos para a constituição do *corpus* a ser submetido à ferramenta computacional para tratamento lexicométrico aos dados, o que permite balizar livremente o texto a ser analisado, determinando e distinguindo as partes do texto previamente digitalizado, além de determinar o tamanho do contexto de um segmento a ser pesquisado e fazer o levantamento das ocorrências do segmento, indicando a distribuição das palavras dentro do texto. A ferramenta computacional permite exibir, ainda, as concordâncias, apresentadas por meio de gráficos, com as frequências relativa e absoluta de uma palavra. O tratamento dos dados oferece, pois, um léxico lematizado que corresponde a um dicionário de construção e uso em que as definições semânticas dos lemas não serão incluídas, mas depreendidas pelos contextos das abonações.

Assim, o verbete constitui-se por: entrada com a norma ortográfica atual, significado(s) no documento, abonação com o local e número do documento, variantes e frequência global. No LHisPB estão contidos registros de alguns dos usos da língua portuguesa com exemplificação dos diversos contextos gramaticais e semânticos extraídos dos manuscritos, datados dos séculos XVII, XVIII e XIX, editados pelas equipes estaduais do PHPB (Paraná, São Paulo, Rio de Janeiro, Bahia, Paraíba e Pernambuco).

1.1. O suporte: apresentação do LHisPB

Ao contrário das demais obras que compõem a série *História do Português Brasileiro*, recentemente publicada no formato impresso, o LHisPB vem à luz no formato eletrônico e disponível para o público em geral na plataforma da Universidade Estadual de Londrina.

A elaboração do LHisPB visa atender, especificamente, aos objetivos de:

- (i) disponibilizar, sob a forma de banco de dados informatizado, um amplo material lexicográfico para fins de pesquisas e estudos sincrônicos e diacrônicos do léxico, com base em documentos manuscritos produzidos, no Brasil, durante os séculos XVII, XVIII e XIX;
- (ii) oferecer aos consulentes um instrumento eficaz de pesquisa lexicográfica quanto ao uso de vocábulos utilizados na escrita de documentos oficiais redigidos e enviados a diferentes destinatários à época do Brasil Colônia e Brasil Império;
- (iii) apresentar as lexias e suas variantes gráficas, juntamente com a frequência de uso e abonações, extraídas de cerca de 1500 documentos manuscritos provenientes de vilas e províncias localizadas nos atuais estados brasileiros.

O LHisPB constitui-se, atualmente, dos seguintes Léxicos estaduais: Léxico Histórico da Paraíba-LHisPAB, Léxico Histórico da Bahia-LHisBA, Léxico Histórico de Pernambuco – LHisPE, Léxico Histórico do Rio de Janeiro-LHisRJ, Léxico Histórico de São Paulo-LHisSP, Léxico Histórico do Paraná-LHisPAR e Léxico Histórico do Rio Grande do Sul-LHisRS. Cada um desses Léxicos foi elaborado com o *corpus* construído a partir dos manuscritos oficiais relativos aos anos 1698 a 1899, produzidos em diversas vilas e províncias do Brasil Pré-Republicano. Dessa forma, todo material que compõe cada um desses Léxicos foi coletado, transcrito e editado pelos pesquisadores das respectivas equipes estaduais e, posteriormente, enviados para a equipe paranaense que lhes deu o tratamento lexicográfico.

Devido ao grande volume e à ampla diversidade de gêneros textuais do material coletado pelas equipes regionais, optamos por trabalhar, majoritariamente, com manuscritos oficiais, na tentativa de compor um *corpus* geral mais homogêneo, cujas entradas e verbetes pudessem ser intercomparados. Dessa forma, o acervo desses documentos oficiais contempla requerimentos, atas, certidões, contratos, declarações, autos, pareceres, relatórios, denúncias, e algumas cartas pessoais no caso dos Léxicos de São Paulo, do Rio de Janeiro e do Rio Grande do Sul, entre outros.

2. Por que a denominação *Léxico* para nosso trabalho?

A responsabilidade, que nos foi dada pelo colega Ataliba Teixeira de Castilho, de compor um Léxico Histórico do Português Brasileiro, implicou desde a sua gênese tomar inúmeras decisões *arriscadas* que expusemos no decorrer deste texto.

Por um lado, sentíamos a urgência e a necessidade de trazer à luz um Léxico dessa natureza, que pudesse auxiliar na complementação de trabalhos já publicados, como os de Cunha (1978, 2006) e Silva (2008).

Acreditamos que o projeto mais ousado na área de dicionários históricos publicados no Brasil seja o idealizado e coordenado inicialmente pela saudosa colega Maria Tereza Biderman. Na ausência de Biderman, Clotilde Muraka assumiu a coordenação e trouxe à luz o *Dicionário histórico do português do Brasil séculos XVI, XVII e XVIII*, em 2021. Sobre a construção do DHPB, Murakawa informa:

Coordenar um projeto de tal magnitude foi tarefa árdua e exigiu a construção de um *corpus* específico, teoria e metodologia lexicográficas adequadas para sua elaboração e busca de documentação referente ao Brasil nos séculos XVI, XVII e XVIII e o 1º decênio do XIX. E onde estavam os documentos e textos necessários para a montagem do *corpus*? Em bibliotecas públicas e particulares, no Brasil e em Portugal, em arquivos, museus e bibliotecas em mosteiros no país. Todo o material recolhido foi digitalizado e inserido no *corpus* que possui 11.133.739 itens e 249.372 formas, permitindo aos redatores elaborarem os 10 470 verbetes que compõem o DHPB, documentando a língua nos 3 séculos da colonização portuguesa no Brasil (MURAKAWA, 2021).

Por outro lado, o Projeto *Para a História do Português Brasileiro* (PHPB) que, desde 1997, reúne pesquisadores de várias IES do Brasil e do exterior em torno da ideia de reconstruir *a história social linguística do Brasil, a partir de fontes históricas múltiplas e comparar o português brasileiro ao europeu, partindo do português de quinhentos em direção aos nossos dias* (MATTOS; SILVA, 1998, p. 10), não poderia prescindir de um trabalho lexicográfico, embora modesto, que marcasse seu 25º aniversário. Esse volume XII – o *Léxico Histórico do Português Brasileiro on-line* – irá se somar aos demais volumes impressos da coleção *História do Português Brasileiro*, publicados sob a direção de Castilho entre os anos 2019 e 2021. Essa coleção trouxe à luz a história de nossa língua em seus aspectos gramatical, lexical, semântico e textual, baseados em fontes organizadas em um amplo *corpus* diacrônico, editado com rigor filológico.

Mas, por que Léxico e não Dicionário ou Vocabulário? Essa é a pergunta que frequentemente nos fazem. A primeira resposta seria: recebemos das mãos do prof. Ataliba Castilho o exemplar do *Léxico Histórico del Español de México* (COMPANY; MELIS, 2002), com o convite-intimação: “Façam o Léxico Histórico do Português Brasileiro”. A leitura da obra mexicana nos seduziu e resolvemos colocá-la em prática, ignorando a amplitude e a exaustão da tarefa.

A segunda resposta, após a decisão de seguir o modelo de Company e Melis, veio da pesquisa em torno da definição de cada trabalho lexicográfico vigente: dicionário, vocabulário, glossário, léxico, entre outros. Cunha e Aguilera (2019) reconhecem que

A tarefa de definir os diferentes tipos de obras lexicográficas não é fácil. A diferença, por exemplo, entre dicionário, léxico, tesouro, vocabulário ou glossário não é consensual entre os estudiosos da matéria. Alguns metalexígrafos se baseiam em critérios linguísticos para estabelecer sua tipologia, outros na história da lexicografia, outros, ainda, em trabalhos lexicográficos já existentes. A convergência de seus métodos está no fato de que quase todos se valem de dicotomias para postular suas definições (CUNHA; AGUILERA, 2019, p. 99).

A tentativa de definir os vários trabalhos lexicográficos não tem sido consensual entre os especialistas. Basta verificar o que dizem alguns lexicógrafos em dicionários dos séculos XIX, XX e XXI.

Moraes Silva (1813), por exemplo, traz “DICCIONARIO, s.m. Vocabulário; livro, que se apontam as palavras de uma Língua e a explicação dos seus significados. LÉXICON, s. m. Diccionario, vocabulários, plur. Léxicos”.

Pinto (1832) registra: “Diccionario s. m. Vocabulário, que contem todas as palavras de huma língua com suas definições. Lexicon s. m. Palavra Grega. He o mesmo que Diccionario”.

Em Caldas Aulete (1958), temos: DICIONÁRIO, s. m. coleção de tôdas as palavras ou somente de certa classe de palavras de uma língua, por ordem alfabética com a sua significação na mesma língua ou com a tradução em outra. LÉXICO, s. m. dicionário de línguas clássicas antigas; dicionário; vocabulário. || Conjunto de vocábulos de uma língua. || Dicionário abreviado. || Vocabulário peculiar a um autor.

O Aulete digital traz quatro definições para DICIONÁRIO: 1. Obra que reúne, em ordem alfabética, as palavras de uma língua ou termos referentes a uma matéria específica, e descreve seu significado, uso, etimologia etc., na mesma língua ou em outra (dicionário de cinema / de inglês). 2. O conjunto das palavras ou termos reunidos nessa obra. 3. Livro ou outro suporte que contém tais informações (dicionário eletrônico). 4. Pessoa de extensos conhecimentos; dicionário ambulante.

Sob o lema LÉXICO, o Aulete digital registra: 1. Ling. Próprio das palavras ou referente a elas; LEXICAL. 2. Ling. O repertório de palavras de uma língua ou de um texto; VOCABULÁRIO. 3. P. ext. Obra de compilação de uma parte (reduzida ou extensa) dos vocábulos de uma língua e seus significados; DICIONÁRIO. 4. Dicionário de antigas línguas clássicas. 5. P. ext. Relação de palavras us. por um autor ou por uma escola ou movimento literário. 6. Ling. Conjunto dos lexemas da língua (proposto por Saussure), oposto ao conjunto de vocábulos. 7. P. ext. Gram. Componente da gramática internalizada de um falante que abarca todo o seu conhecimento das palavras (esp. sua pronúncia, significação e emprego numa sentença).

Cientes do dissenso entre os especialistas na definição de *dicionário* e *léxico* e diante do fato de: (i) termos selecionado textos da mesma natureza (manuscritos oficiais sincrônicos); (ii) não termos incluído a definição de cada lema, mas (iii) termos documentado exaustivamente o emprego de cada variante gráfica ou morfológica, consideramos legítima a nomenclatura LÉXICO atribuída ao *Léxico Histórico do Português Brasileiro*.

3. Procedimentos metodológicos: da obtenção dos dados à acessibilidade ao *site*

Nesta seção, considerando que cada um dos *corpora* estaduais foi composto de acordo com o interesse das respectivas equipes, atendendo também às orientações da coordenação geral, discorreremos sobre os procedimentos metodológicos levados a efeito na elaboração dos Léxicos estaduais, tratando: (i) da descrição dos documentos estaduais e da abrangência temporal; (ii) do tratamento lexicográfico dado pela equipe paranaense; (iii) da preparação e publicação do *site* e o acesso a ele.

3.1. Descrição dos documentos que compõem o *corpus* do LHisPB

Os documentos que compõem os vários Léxicos estaduais que integram o LHisPB procedem, normalmente, de arquivos públicos, no formato original, ou em *fac-símiles* e, no caso do Léxico do Rio Grande do Sul, de obra já editada por outrem.

O recorte temporal e o local de depósito, ou origem do material, são diversos conforme demonstramos no quadro 1.

Quadro 1: Fonte dos Léxicos Estaduais segundo a datação dos documentos e o respectivo local de depósito:

Estado	Tipologia dos Documentos	Datação	Origem/depósito
PR	Documentos oficiais	1731 - 1853	Arquivo Público de SP-UEL
SP	Documentos oficiais, Processos crimes e Cartas pessoais	1796 - 1890	Arquivo do Estado de São Paulo, Tribunal de Justiça e Museu Paulista-USP
RJ	Documentos oficiais e Cartas pessoais	1720 - 1896	www.lettras.ufrj.br/phpb-rj-UFRJ
BA	Documentos oficiais	1698 - 1828	PROHPOR-UEFS
PB	Documentos oficiais	1774 - 1874	Maria Cristina Fonseca (coord.) UFPB
PE	Documentos oficiais	1704 - 1874	Cleber Ataíde e Ana Paula Macena (coords.) UFPE
RS	Documentos oficiais e Cartas pessoais	1835 - 1845	Anais do Arquivo Histórico do RS-Arquivo Histórico do RS

Fonte: Elaborado pelas autoras com dados da pesquisa.

O *corpus* do Paraná consta de 728 fólhos referentes a 437 manuscritos oficiais redigidos entre os séculos XVIII e XIX, mais especificamente de novembro de 1721 a dezembro de 1853, no âmbito das antigas vilas, como as de Paranaguá, Curitiba, Guaratuba, Antonina, Morretes, Iapó (atualmente Castro), Vila Nova do Príncipe (hoje Lapa), São José, entre outras. As cópias digitalizadas desses manuscritos, disponíveis no *site* do Projeto do Léxico Histórico do Paraná (LHisPar), foram obtidas junto ao Arquivo Público do Estado de São Paulo, uma vez que o território atual do Paraná pertenceu à Província de São Paulo até 20 de dezembro de 1853, quando se deu a sua emancipação. O *corpus* paranaense foi editado de acordo com as normas prescritas pelo PHPB (MEGALE, 2001, pp. 553-5), com adaptações pertinentes à proposta da equipe de autores da edição desses manuscritos.

Para compor o *corpus* de São Paulo, buscamos: (i) Edição semidiplomática de *Memória Histórica da Capitania de São Paulo e todos os seus Memoráveis Sucessos desde o anno de 1531 até o presente 1796*, (ii) sete processos-crime depositados no Arquivo do Tribunal de Justiça do Estado de São Paulo e (ii) 28 cartas pessoais, do Museu Paulista, referentes às Correspondências particulares do Fundo Rafael Tobias de Aguiar, editadas por Alexandra de Souza Marques, sob a supervisão de Marcelo Módolo, da Universidade de São Paulo.

Sobre o *Memória Histórica*, Costa (2014, p.11), esclarece:

Pesquisando o livro de registros do Arquivo, o título de um documento chamou a atenção: *Memória Histórica da Capitania de São Paulo e todos os seus memoráveis sucessos desde o anno de 1531 até o presente de 1796*, por situar-se justamente em fins do século XVIII e compreender dados da história de São Paulo, cidade atualmente reconhecida como importante núcleo de atividades intelectuais, políticas e econômicas, mas que até o século XVIII ocupou uma posição quase insignificante na colônia portuguesa, situação acarretada pela distância do litoral, pelo isolamento comercial e pela carência de uma atividade econômica lucrativa.

Ao todo, o *corpus* do LHisSP consta de 27 316 palavras que elencaram 5357 formas para a elaboração dos verbetes.

O acervo manuscrito utilizado para a elaboração do Léxico Histórico do Rio de Janeiro⁴ compõe-se de 141 cartas de natureza pessoal e oficial, datadas dos séculos XVIII e XIX, a primeira das quais de 1720 e a última de 1896. O *corpus* do LHisRJ composto de documentos da Administração Pública, documentos da Administração Privada e documentos particulares enviados a Rui Barbosa contém 54 397 palavras e 9023 formas. Não dispomos dos *fac-símiles* uma vez que os documentos estão disponíveis no modelo de edição diplomático-interpretativa⁵ realizada por Barbosa (1999), conforme consta de sua tese. Trata-se de documentação oficial e de cartas de comércio do Brasil no final século XVIII.

Os documentos integram o acervo da Biblioteca Nacional do Rio de Janeiro e provêm de coleções como a de Morgado de Mateus, Marquês do Lavradio e cartas pessoais endereçadas ao então Conselheiro Rui Barbosa, entre outros.

Barbosa (1999, p. 308-310), em sua tese, ao expor as Normas para Transcrição dos Manuscritos do Século XVIII, inicia alertando que:

a transcrição da grafia dos textos reunidos em nosso *corpus* é de inspiração conservadora. Não se busca, em contrapartida, uma reprodução fac-similada, mas, na verdade, um ponto de equilíbrio entre o desejo de indicar hábitos gráficos da época e o de oferecer um texto de fácil leitura para o leitor contemporâneo não-especialista.

Da Bahia, para integrar o LHisPB, buscamos os documentos oficiais, inventários, testamentos e declarações que constam do projeto coordenado por Ilza Ribeiro, Zenaide de Oliveira Novais Carneiro e Norma Lucia Fernandes de Almeida, *Contribuição para a constituição de um banco de textos e de um banco de dados para o estudo da história do português no Brasil*, do século XVII ao XIX. Datados entre os anos de 1698 e 1828, os manuscritos foram editados por Novais e Carneiro

⁴ O LHisPB serviu-se dos documentos do Rio de Janeiro alocados no *site* www.lettras.ufrj.br/phpb-rj, que, atualmente está no <https://sites.google.com/site/corporaphpb/home/corpora-manuscritos/tabela-manuscritos-rio-de-janeiro>.

⁵ Os critérios aqui empregados resultam da combinação entre a orientação geral do Professor Doutor Ivo Castro e as — Normas para transcrição de documentos manuscritos definidas por comissão especial sob a presidência do Professor Doutor Heitor Megale do Projeto Para a História do Português Brasileiro (BARBOSA, 1999, p. 308, nota de rodapé 337).

e apresentados em CD-ROM. O *corpus* integra o PROHPOR - Programa para a História da Língua Portuguesa Instituto de Letras, UFBA, durante muitos anos coordenado pela saudosa Rosa Virgínia Mattos e Silva. Ao todo, são 17 877 palavras que deram origem a 3354 formas para a construção dos verbetes.

O *corpus* da Paraíba faz parte do acervo do Projeto História do Português da Paraíba (PHPPB), coordenado por Dermeval da Hora. Para o LHisPB utilizamos 203 cartas oficiais lavradas durante os séculos XVIII e XIX, que foram editadas e organizadas por Fonseca (2004). A primeira carta data de abril de 1774, a penúltima, de agosto de 1874 e da última não constam data nem local. Quanto ao local de origem das cartas, 44 delas partiram de Pernambuco, fato explicável pelo pertencimento da Paraíba àquele estado até o ano de 1799. O *corpus* do LHisPB compõe-se de 37 160 palavras que renderam 6337 formas.

Os documentos de Pernambuco integram o acervo do Conselho Histórico Ultramarino – Conselho Ultramarino Brasil-Pernambuco e o *corpus* do LHisPE contém 18 693 palavras que representam 3939 formas. Coletados pela equipe regional do PHPB-PE, foram editados por Iapechino, da Universidade Federal Rural de Pernambuco e Athayde, da Universidade Federal de Pernambuco. São 21 cartas oficiais do Século XVIII, datadas de 1704 a 1784; 21 cartas oficiais do século XIX, escritas entre 1801-1827; as demais cartas foram escritas entre 1884-1891 pelo Deão Joaquim Francisco de Faria e pelo Cônego Antônio Arcoverde d’Albuquerque Cavalcanti, do Gymnasio Pernambucano, tratando de assuntos referentes a esse estabelecimento de ensino.

Para a elaboração do Léxico Histórico do Rio Grande do Sul, considerando que na época não havia uma equipe estadual integrada ao PHPB, recorreremos à obra *Uma República contra o Império: A Coleção Varela – Documentos sobre a Revolução Farroupilha – 1835/1845*, que consta dos Anais do Arquivo Histórico do Rio Grande do Sul – volume 16), Coleção Alfredo Varela; correspondências de Bernardo Pires e outros. Por se tratar de obra resultante de transcrição de manuscritos⁶ feita por historiadores⁷ e graduandos do curso de História da PUC-RS, e o tema ser de interesse para a história do Português Brasileiro, seu conteúdo foi incluído no LHisRS. O *corpus* consta de 75 688 palavras e 8786 formas.

Os *corpora* de todos os estados, exceto o *corpus* do Paraná e o do Rio Grande do Sul, foram editados pelas equipes regionais integrantes do PHPB. Desse modo, a equipe do Paraná incumbiu-se exclusivamente da edição dos manuscritos paranaenses e de fazer alguns ajustes na edição dos documentos de Pernambuco.

⁶ Transcrição e pesquisa de Paulo Staudt Moreira, Historiógrafo do AHRS; Rejane Penna, Historiógrafa do AHRS, Aline Souza Alves, Estagiária / Graduanda História – FAPA); Diogo de Moraes Vieira, Estagiário / Graduando História – ULBRA (2008, p. 6).

⁷ Sobre as normas seguidas para a edição dos manuscritos, veja a Apresentação Técnica (2008, p.11).

3.2. Procedimentos metodológicos: do tratamento dos dados e da composição dos verbetes

Uma vez recebidos os documentos das equipes estaduais, procedemos aos seguintes passos para a composição do respectivo Léxico Histórico: (i) leitura dos documentos; (ii) limpeza e identificação de palavras homógrafas; (iii) gravação do *corpus* em caixa alta e preparação para submeter à ferramenta Léxico 3 no formato *.txt; (iv) elaboração de planilha com a sequência dos documentos de cada equipe (número, data, procedência, assunto, autor); (v) elaboração dos verbetes segundo a sequência apresentada no Léxico 3; (vi) revisão do Léxico; e (vi) inserção no *site*.

O quadro 2 mostra, em síntese, os dados atuais do Léxico Histórico do Português Brasileiro (LHisPB).

Quadro2: Descrição dos *corpora* estaduais que compõem o *corpus* do LHisPB:

Léxico/Estado	Nº Documentos	Nº de Palavras	Nº de Formas	Nº Verbetes
LHisPaR	37	132 380	14 254	3 263
LHisSP	35	27 316	5 357	2 610
LHisRJ	131	54 397	9 023	3 502
LHisBA	24	17 877	3 354	1 329
LHisPaB	03	37 160	6 337	2 887
LHisPE	64	18 693	3 939	1 797
LHisRS	332	75 688	8 786	3 515
Total	1 523	363 510	51 050	18 903

Fonte: Elaboração das autoras com base nos *corpora* estaduais.

O quadro 2 mostra que, dentre os *corpora* coletados pelas equipes de pesquisadores estaduais, as coordenadoras do LHisPB aproveitaram entre 24 e 437 documentos de cada projeto estadual para compor os verbetes do respectivo Léxico. Essa disparidade numérica pode ser explicada pelo seguinte: (i) por se tratar de documentos paranaenses, o LHisPaR é o mais extenso de todos eles e foi o primeiro *corpus* a ser trabalhado pela equipe. Os documentos se estendem de 1721 a 1828 e compreendem 734 fôlios, correspondendo a 437 documentos; (ii) o LHisRS é o segundo em número de documentos (332) referentes ao período de 1835 a 1845, escritos durante a Revolução Farroupilha; (iii) o número de *formas* corresponde ao levantamento apresentado pela ferramenta *Lexico3*, *software* que auxilia na análise da frequência, contagem, agrupamento dos vocábulos e localização das abonações usadas nos verbetes. Trata-se de programa desenvolvido pela equipe YLED-CLA2T, da Universidade de Sorbonne Nouvelle-Paris 3.

Para a construção do verbete, o conjunto dos documentos transcritos pelas equipes estaduais foi convertido em blocos de textos e aplicados à ferramenta *Léxico3*. A lista gerada por esse programa, além de facilitar a leitura e a compreensão dos termos registrados nos documentos manuscritos e apresentar, de modo geral de ‘A a Z’, todos os vocábulos e suas respectivas variantes lexicais, fornece também dados linguísticos que se constituem num significativo *corpus* diacrônico, acessível a pesquisadores interessados em estudar variantes lexicais de documentos setecentistas, oitocentistas e novecentistas produzidos em vilas e províncias sediadas nos estados do Brasil. Assim, cada um dos *corpora* no formato *.txt foi submetido à ferramenta *Léxico 3* que traz, no lado esquerdo, as formas em ordem lexicográfica (ordem alfabética crescente) ou em ordem lexicométrica (pelo número de ocorrências no texto); no lado direito, o texto do *Léxico* estadual em caixa alta.

A figura 1 traz uma amostra da ferramenta *Léxico3* que serviu de base para o tratamento do *corpus* escrito.

Figura 1: Exemplo de tratamento do *corpus* escrito: listagem dos vocábulos em ordem alfabética, número de frequência e texto transcrito para a submissão ao *Léxico3*

Formes (ordre lexicographique)	Fréquence
a	478
á	1
abaiaru	1
abaino	1
abastados	1
abesta	1
abestas	1
abestos	1
abertura	2
abégivel	1
abitadores	1
abrangem	1
abrigo	1
abú	6
abrio	1
abris	3
abrissem	1
absurdo	1
absurdos	1
abundancias	1
acabada	1
acabe	1
acabou	2
acautelar	1
acautela	1
aceite	1
acenos	1
acentarã	1
acentavam	1
acha	14
acham	1
achandoce	2
achaõ	2
achar	3
acharem	2
achassem	1

excelentissimo senhor
aos de outubro recebemos a carta
de vossa excellencia com a qual nos fazia a honrra
e muita de notesiarnos a decisaõ que sua magestade
que deos guarde foi seruido fazer das capitancias
de sam paullo e agregar esta villa
ao governo de vossa excellencia hem se deixa ver
a real benignidade de sua magestade em
quierer faturecer e amparar a estes
seus vassallos que pella distancia
dos senhores generais naõ podiaõ ser
tambem ouuidos nos seus requerimentos
como o serem agora de vossa excellencia como sua
ihstre fidalgia nos está asegurado
registouçe a carta como vossa excellencia ordena
no livro dos registos e consta da çertidam
que apresentamos
como os offçiais desta camera estiuçem
retirados por suas fazendas pera
benefiço de suas lauouras vnico trato
desta villa naõ nos foi poçiuel com mais
breuidade o podermos dar comprimento
a ordem de vossa excellencia em enviarmos hum
dos republicanos desta villa e como neste
senado os juizes se achem com o procurador
tam ocupados na arrecadaçaõ dos bens
deste senado e o tempo prosimo a darem
contas aos suçesores e o vereador
primeiro com a inuicibilidade de oitenta

Fonte: Lexico3 (<http://www.tal.univ-paris3.fr/lexico/lex3-10pas/index.htm>)

Para verificar o contexto da palavra a ser inserida como entrada no *Léxico Histórico*, a ferramenta *Léxico 3* oferece a opção *Concordances et inventaires distributionels* (3º ícone, na parte superior), centralizando em azul o item buscado, conforme ilustramos com a figura 2.

Figura 2: Exemplo de tratamento do *corpus* escrito: frequência do vocábulo *ballandra* e suas ocorrências no documento manuscrito para busca das contextualizações no LHisPB.

The screenshot shows the Lexico3 software interface. On the left, there is a search results table with the following data:

Formes (ordre lexicographique)	Fréquence
avaliar	1
avaliados	1
avaliar	2
avaliar	1
avaliado	1
avencia	1
avanzado	1
aver	5
avermos	1
avia	3
aviam	1
avistado	1
avocar	1
avulsam	1
avão	1
az	1
áz	1
azamor	4
azeite	1
baña	1
balhas	1
baño	2
banos	1
ballandra	2
baltear	1
balhazar	1
bando	7
baptizados	1
barbosa	1
barboza	5
barcos	2
bara	4
barat	2
barrete	1
barrette	1
barstite	4

The main text view on the right shows a snippet from a manuscript: "o nesta villa chegou aos dois do março passado huã ballandra enviada pelo governador da nova collonia com hu- cabo por nos mandado apenar o povo para a tara de que a ballandra carecese e com effeito envidias recebo o que pode".

Fonte: Lexico3 (<http://www.tal.univ-paris3.fr/lexico/lex3-10pas/index.htm>)

Para a construção dos verbetes, utilizamos a tabulação das entradas fornecida pela frequência dos dados extraídos do *Léxico3* e, em seguida, aplicamos os mesmos critérios metodológicos adotados por Company e Melis (2002).

4. O *site* do LHisPB: construção, descrição, publicação e acesso

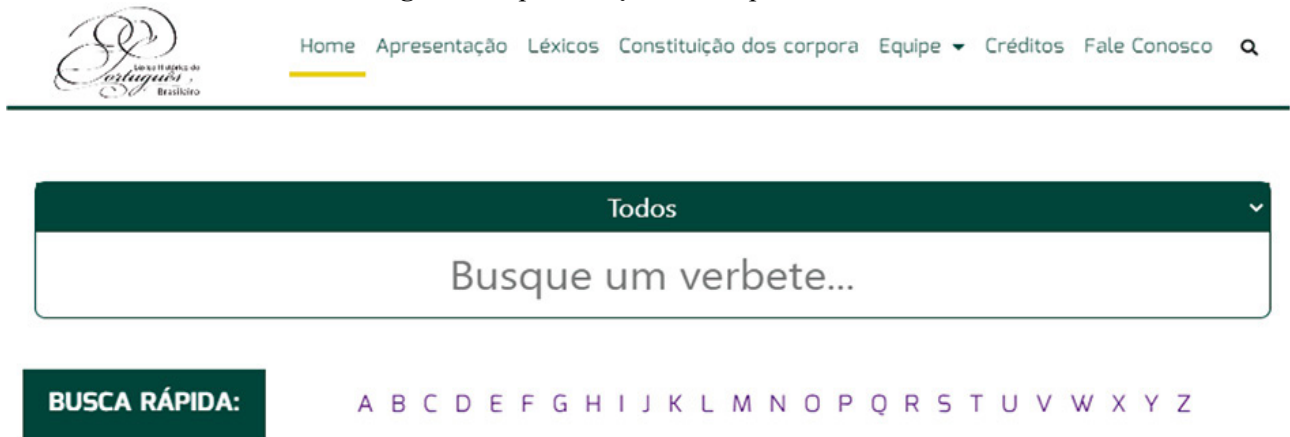
Concluídas algumas etapas de trabalho para a elaboração do Léxico Histórico do Português Brasileiro, procedemos à construção dos verbetes seguindo os moldes da obra de Company e Melis (2002): (i) entrada com a norma ortográfica atual, (ii) contextos que melhor exemplificam o emprego sintático de todas as formas gramaticais registradas no *corpus*-base, (iii) abonação com o local e datação e (iv) número do documento, variantes e frequência global.

Cumprida essa etapa, dedicamo-nos à criação do *site* do LHisPB⁸ com o intuito de oferecer um banco de dados informatizado mediante uma consulta rápida dos vocábulos extraídos de manuscritos datados dos séculos XVII, XVIII e XIX editados pelas equipes estaduais do PHPB.

⁸ O *site* encontra-se em revisão dos dados e ajustes internos e pode ser acessado pelo link <https://www.uel.br/lhispb/>.

A apresentação do *site* para o LHisPB estrutura-se conforme a ilustração (figura3).

Figura 3: Apresentação do *site* para o LHisPB



Fonte: (<http://www.tal.univ-paris3.fr/lexico/lex3-10pas/index.htm>)

Na página principal (*home*) constam as seguintes informações: (i) logomarca que dá acesso à página principal, (ii) texto de apresentação, (iii) léxicos, (iv) constituição dos *corpora*, (v) equipe, (vi) créditos, e (vii) fale conosco. Na aba *home* também é possível fazer uma busca rápida digitando o verbete desejado no campo de pesquisa, como também o consulente pode clicar em cima de qualquer letra alfabética e pesquisar o vocábulo que consta do banco de dados do *corpus* como um todo, ou filtrar para o Léxico Histórico do Português de uma das sete regionais catalogadas, a saber: Rio Grande do Sul (LHisRS), Rio de Janeiro (LHisRJ), Paraná (LHisPAR), São Paulo (LHisSP), Pernambuco (LHisPE), Paraíba (LHisPB) e Bahia (LHisBA). No campo para exibição dos resultados é possível encontrar: a abonação do verbete, número de variantes documentadas, frequência global, imagem do *fac-símile*, transcrição do manuscrito e o mapa com a localização da cidade no qual pertence o documento.

A título de ilustração, nas figuras 4 e 5, trazemos o resultado da busca do verbete *propina* no Léxico Histórico do Paraná (LHisPAR).

Figura 4: Exemplo de busca por verbete no do LHisPB.

propina

Proveu que se obcervasse, e executasse as provizoins de vinte, e dous de março de mil oitocentos, e tres expedidas pelo Conselho ultramarino relativamente a *propina* que esta camera deve satisfazer anualmente ao secretario do mesmo tribunal (PGA. 1804. 498. 16): Attendido por nós a de Vossa Excelencia nos juntamos com o Doutor Antonio dos Santos Soares, ouvidor desta comarca; para effeito de se arbitrar o quanto poderia dar esta camera de *propinas* ao secretario achou o dito ouvidor (PGA. 1732. 623. 10).

Formas Documentadas

propina (5), propinas (6).

Frequência Global

11

Registros em:

Paranaguá

Transcrição

[[30-Dez-1] [[Paranaguá]] [[12-422]] Exceletissimo Senhor [[1732]]

Aas 17 de Outubro recebemos e veneramos as de vossa Excelencia de 12 de Setembro, e com ella o bando que vossa Excelencia foi servido mandar se publicou o que logo foi satisfeito como da inclusa certidã vossa Excelencia verá.

Attendido por nós a de Vossa Excelencia nos juntamos com o Doutor Antonio dos Santos Soares, Ouvidor desta comarca; para effeito de se arbitrar o quanto poderia dar esta Camera de propinas ao Secretario achou o dito Ouvidor; attendido por elle, e por nos as pocibilidades desta Camera; arbitrar 4 reis que consta da certidã junta, que vaj com o theor do termo do arbitramento a qual fica a esta a ordem de Vossa Excelencia detriminarã se se lhe hade Remeter ou entregar ca a quem o fasa.

Tambem vemos o Paternal afeto com que Vossa Excelencia se offerece attender aos particulares deste Senado, e Representallos a Real grandeza; no que vive este Povo gratuita mente confiado, e assim nós ficamos Rogando a Deus prospere a vida e saude a Vossa Excelencia para amparo desta cappitania Deos guarde a Vossa Excelencia muitos annos Paranaguã de Dezembro de 1732

De Vossa Excelencia
dos Sarzedas
Os mais humildes subditos
Teixeira Marinho Joãõ Davaeiga Dasiqueira
Luis Matozo Pedro Reis

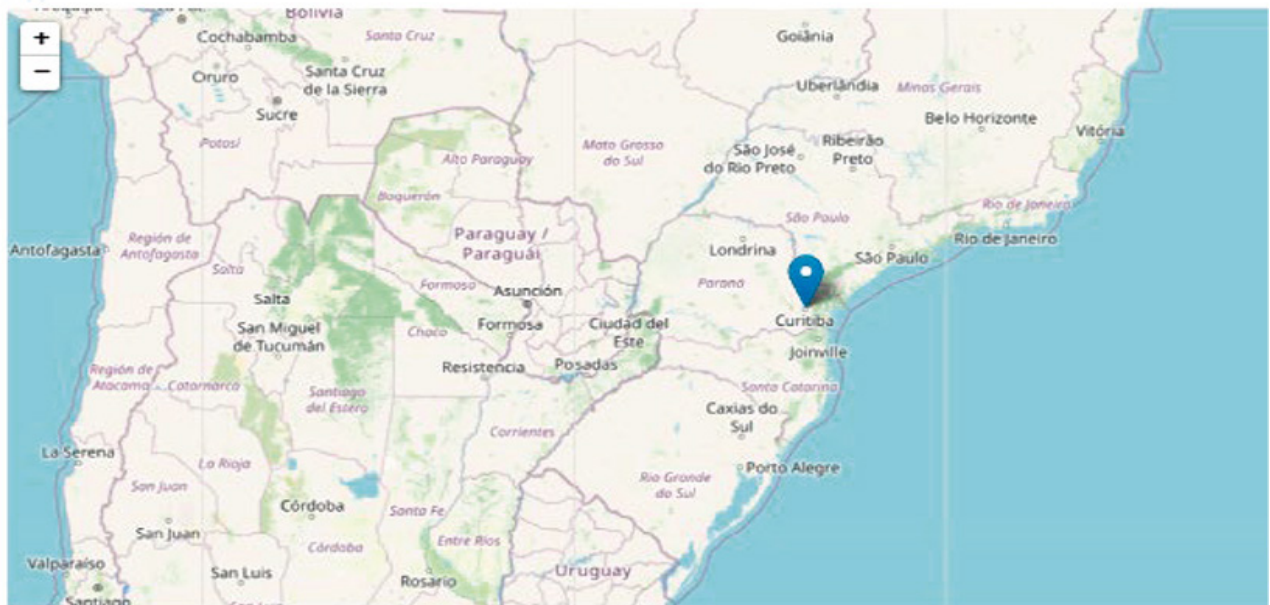
Manuscrito:



Fonte: Lexico3 (<http://www.tal.univ-paris3.fr/lexico/lex3-10pas/index.htm>)

Figura 5: Exemplo de busca pelo local de origem do verbete no do LHisPB

Mapa



Fonte: Lexico3 (<http://www.tal.univ-paris3.fr/lexico/lex3-10pas/index.htm>)

À guisa de uma conclusão

Ao longo desses anos, de sua ideia inicial ao momento de consolidação da apresentação dos dados das equipes do PHPB, o LHisPB tem movimentado esforços no sentido de propiciar a acessibilidade a todos aos dados coletados. Os procedimentos metodológicos para o tratamento dos verbetes vêm assegurar a preservação dos dados coletados, permitindo o acesso da comunidade científica por meio do léxico de frequência. Salientamos a importância da organização do Banco de Dados, disponibilizado pelo LHisPB, instrumento de pesquisa que permite a documentação linguística sistematizada, tornando-se objeto de estudo de diversas áreas da Linguística.

Destacamos, ainda, a importância do acesso ao acervo linguístico regional: o Léxico, importante veículo para a conservação e disseminação do acervo linguístico coletado, possibilitará a acessibilidade à síntese dos dados coletados. Muitos estudiosos poderão se servir dele para suas pesquisas, permitindo o entendimento dos usos linguísticos, dentro do contexto social da comunidade estudada.

Estão todos convidados para um passeio pelo *site*!

Referências

AGUILERA Vanderci de Andrade; ALTINO Fabiane Cristina. Léxico histórico do português brasileiro. In: HORA D, Silva CR. *Para a história do Português Brasileiro: abordagens e perspectivas*. João Pessoa: Ideia/ Editora Universitária, 2010, pp. 1-10.

ALMEIDA-BARONAS, Joyce Elaine. Manuscritos paranaenses: fontes para um estudo diacrônico. In: *Anais do V SELISIGNO e VI Simpósio de Leitura* da Universidade Estadual de Londrina, Londrina – PR, 2006.

ALMEIDA-BARONAS, Joyce Elaine; ALMEIDA, Polyana Lucena Camargo de. *A escrita dos séculos XVIII e XIX: em busca de dados diacrônicos*. In: *XIX Seminário do CELLIP – Centro de Estudos Linguísticos e Literários do Paraná*. UNIOESTE – Universidade Estadual do Oeste do Paraná, Cascavel – PR, 2009.

ALTINO, Fabiane Cristina; AGUILERA, Vanderci de Andrade. Léxico Histórico do Português Brasileiro. In: HORA, Dermeval da; SILVA, Camilo Rosa (org.). *Para a História do Português Brasileiro: abordagens e perspectivas*. João Pessoa: Ideia/ Editora Universitária, 2010.

ARQUIVO HISTÓRICO DO RIO GRANDE DO SUL Uma República contra o Império [recurso eletrônico]: (1835/1845): *Anais / Arquivo Histórico do Rio Grande do Sul*. – Dados eletrônicos – Porto Alegre: EDIPUCRS, 2009 – (Coleção Varela, v. 16).

AULETE Caldas. *Dicionário contemporâneo da língua portuguesa*. 4. ed. Rio de Janeiro: Delta, 1958.

AULETE Digital. *Dicionário contemporâneo da língua portuguesa: Dicionário Caldas Aulete, on-line*. Disponível em: <https://www.aulete.com.br/online>. Acesso em: 12 abr. 2023.

BARBOSA, Afrânio Gonçalves. *Para uma história do Português Colonial: aspectos linguísticos em Cartas de Comércio*. Tese (Doutorado em Letras). Rio de Janeiro: Universidade Federal do Rio de Janeiro, 1999.

- BENÇAL, Dayme Rosane. *Registros escritos paranaenses dos séculos XVIII e XIX: um estudo sobre o acento das médias pretônicas <e>*. 2018. 253f. Tese (Doutorado em Estudos da Linguagem). Londrina: Universidade Estadual de Londrina, 2018.
- BIDERMAN, Maria Teresa Camargo (*in memoriam*); MURAKAWA, Clotilde de Almeida Azevedo (org.). *Dicionário Histórico do Português do Brasil: séculos XVI, XVII e XVIII*. 1. ed. Araraquara - SP, v. 1, 2021.
- COMPANY, Concepción; MELIS, Chantal. *Léxico histórico del español de México*. México: Universidad Autónoma de México, 2002.
- COSTA, Renata Ferreira. *Memória histórica da Capitania de São Paulo: edição e estudo*. São Paulo: Arquivo Público do Estado de São Paulo, 2014.
- CUNHA, Antônio Geraldo da. *Dicionário Etimológico da Língua Portuguesa*. s/l: Lexikon, 2006.
- CUNHA, Antônio Geraldo da. *Dicionário histórico das palavras portuguesas de origem tupi*. 5. ed. São Paulo: Companhia Melhoramentos; Brasília: Universidade de Brasília, 1978.
- CUNHA, Cláudio de Assis da. *Edição semidiplomática de manuscritos de Antonina: traçando caminhos da filologia*. 216f. Dissertação (Mestrado em Estudos da Linguagem) – Universidade Estadual de Londrina, Londrina, 2008.
- CUNHA, Cláudio de Assis da. *O léxico em documentos produzidos no Paraná - Brasil (Séculos XVIII e XIX): Abordagem lexicográfica*. 239f. Tese (Doutorado em Estudos da Linguagem) – Universidade Estadual de Londrina, Londrina, 2013.
- CUNHA, Cláudio de Assis da; AGUILERA, Vanderci de Andrade. Tipologia das obras lexicográficas e o Léxico Histórico do Português Brasileiro. *Filologia Linguística Portuguesa*. São Paulo, v. 21, n. 1, pp. 99-114, 2019.
- DHPB - *Dicionário Histórico do Português do Brasil* - <https://www.fclar.unesp.br/#!/noticia/1204/lancamento-do-dicionario-historico-do-portugues-do-brasil/>. 2021. Acesso em: 12 abr. 2023.
- FONSECA, Maria Cristina de Assis Pinto (org.). *História do Português da Paraíba (PHPPB)*. João Pessoa: Idéia. 2004.
- IAPECHINO, Mari Noeli Kiehl. *Cartas Oficiais* – Pernambuco. Recife: Projeto PHPB /PE, 2010, CD-rom. Cartas Oficiais - Carta 1.
- MEGALE, Heitor. *et al.* Normas para transcrição de documentos manuscritos para a História do Português do Brasil. In: Mattos e Silva, Rosa Virgínia. *Para a História do Português Brasileiro*. VOL. II: Primeiros Estudos, Tomo II. São Paulo: Humanitas/FFLCH/USP: FAPESPA, 2001, pp. 553-5.
- MORAES SILVA, Antonio de. *Diccionario da lingua portugueza recompilado dos vocabularios impressos ate agora, e nesta segunda edição novamente emendado e muito acrescentado, por Moraes Silva A.* Lisboa: Typographia Lacérdina; 1823. Disponível em: <https://digital.bbm.usp.br/handle/bbm/5413>.
- PINTO, Luiz Maria da Silva. *Diccionario da lingua brasileira*. Ouro Preto: Typographia de Silva; 1832. [citado 20 out. 2010]. Disponível em: <https://digital.bbm.usp.br/handle/bbm/5414>.

SILVA, Joaquim Carvalho da. *Dicionário da Língua Portuguesa Medieval*. Londrina – PR: EDUEL, 2008.

SOUZA, Adriana dos Santos. *Contribuições para o estudo da gramaticalização do item ONDE na história do Português Brasileiro: documentos paranaenses*. 2012. 193f. Tese (Doutorado em Estudos da Linguagem) – Universidade Estadual de Londrina, Londrina, 2012.

TONIOLO, Enio. *Possíveis marcas da oralidade em manuscritos paranaenses*. In: *Filologia e Linguística Portuguesa*. FLP 10/11, 2008. Disponível em: www.fflch.usp.br/dlcv/lport/flp/images/arquivos/FLP10-11/Toniolo.pdf.