

Para a história do português popular brasileiro: edições fac-similar, semidiplomática e modernizada de Cartas Marienses (século XX)

For the history of popular Brazilian portuguese: fac-similar, semidiplomatic and modernized editions of Cartas Marienses (20th Century)

Patrícia Santos de Jesus Brito*
Universidade Estadual de Feira de Santana, BA, Brasil

Mariana Fagundes de Oliveira Lacerda**
Universidade Estadual de Feira de Santana, BA, Brasil

Resumo: Este trabalho apresenta o acervo *Cartas Marienses*, do banco de dados do projeto *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). O referido acervo soma-se à documentação epistolar do CE-DOHS, reunida para atender a uma das agendas de pesquisas do *Projeto Nacional para a História do Português Brasileiro* (PHPB). As *Cartas de Coração de Maria* (Bahia), datadas do século XX, representam as normas socialmente estigmatizadas do PB; trata-se de um dos três acervos do banco em questão, com base nos quais se pode contar uma história do português popular brasileiro, no interior da Bahia. No presente texto, apresentam-se as edições fac-similar, semidiplomática e modernizada das *Cartas Marienses*, abordando a importância da constituição de banco de dados para uma história social e linguística do PB, com destaque para as fontes do português popular brasileiro (Mattos e Silva, 2004).

Palavras-chave: Português popular brasileiro. Banco de dados. Sertão Baiano.

Abstract: This study presents the collection *Cartas Marienses*, from the database of the project *Corpus Eletrônico de Documentos Históricos do Sertão* (CE-DOHS), which is part of the center of study *Núcleo de Estudos de Língua Portuguesa* (NELP) of the Universidade Estadual de Feira de Santana (UEFS). The mentioned collection composes part of the epistolary documentation of the project CE-DOHS, gathered to attend one of the research agendas of the national project *Para a História do Português Brasileiro* (PHPB). The collection of letters *Cartas de Coração de Maria* (Bahia), dated in 20th century, represent the stigmatized social norms of the Brazilian Portuguese (PB); it is one of three collections of the data bank, and based on them it is possible to tell a history of non standard Brazilian Portuguese, in the interior of Bahia. In this work, it is presented the fac-similar, semi-diplomatic and modernized edition of the *Cartas Marienses* collection, addressing the importance of structuring a database for a social and linguistic history of PB, emphasizing on the sources of non standard Brazilian Portuguese.

Keywords: Non standard Brazilian Portuguese. Database. Sertão of Bahia.

* Doutoranda pelo Programa de Pós Graduação em Estudos Linguísticos (PPGEL) da Universidade Estadual de Feira de Santana, BA, Brasil; patysantosjesus@hotmail.com

** Professora Titular do Departamento de Letras e Artes da Universidade Estadual de Feira de Santana, BA, Brasil; marianafag@gmail.com

1 CONSIDERAÇÕES INICIAIS

A reconstrução da sócio-história linguística do português brasileiro (PB) exige fundamentalmente a reconstrução da história social linguística – a partir de uma sociolinguística histórica ou sócio-história linguística – e de uma história linguística do Brasil (Mattos e Silva, 2004). Para isso, fazem-se necessárias fontes históricas múltiplas (Mattos e Silva, 2002), as quais podem aclarar o percurso histórico do PB. Essas fontes históricas, testemunho linguístico daqueles que fizeram o PB, se multiplicam, segundo Mattos e Silva (2004), desde 1500, com a carta de Caminha.

O *Projeto para a História do Português Brasileiro* (PHPB), desde sua criação, em 1997, busca reconstruir o passado do PB ao desenvolver trabalhos, os quais cumprem as seguintes agendas de pesquisas, conforme descritas por Lobo (2009): (a) constituição de *corpora* diacrônicos de documentos de natureza vária, escritos no Brasil, a partir do século XVI; (b) a reconstrução da história social linguística do Brasil; e (c) estudos de mudanças linguísticas depreendidas na análise dos *corpora* constituídos.

No Núcleo de Estudos de Língua Portuguesa (NELP), da Universidade Estadual de Feira de Santana (UEFS), pesquisadores do projeto *Corpus Eletrônico de Documentos Históricos do Sertão* (CE-DOHS) realizam as edições fac-similar, semidiplomática e modernizada de textos – especialmente documentação epistolar –, segundo critérios de transcrição do PHPB (Mattos e Silva, 2001), do qual o CE-DOHS é parceiro, para estudo de diversos aspectos linguísticos, com análises desenvolvidas dentro de diferentes quadros teóricos.

Neste trabalho, apresentam-se as edições fac-similar, semidiplomática e modernizada do acervo *Cartas Marienses*, que consiste em um conjunto de cartas, de caráter pessoal, escritas no decorrer do século XX, por grupos socialmente estigmatizados, da região rural de Coração de Maria, interior baiano. Essas edições pretendem amenizar a lacuna que existe no que diz respeito às fontes históricas representativas do português popular brasileiro. Como destaca Santiago (2012), há uma quantidade generosa de fontes históricas representativas do português culto brasileiro; no entanto, em relação à vertente popular, há uma escassez de fontes, sobretudo porque esse português se constituiu na oralidade.

O trabalho em questão está organizado da seguinte forma: em 1, as considerações iniciais; em 2, discorre-se sobre a constituição do banco de dados para a história social e linguística do Brasil; em 3, fazem-se algumas pontuações sobre o banco de dados do NELP/UEFS; em 4, discute-se sobre o PB popular e sobre as fontes de estudos; em 5, apresenta-se a contextualização sócio-histórica do corpus *Cartas Marienses*; em 6, apresentam-se as edições fac-similar e semidiplomática; e em 6.1, apresenta-se a edição modernizada; seguida das considerações finais.

2 A CONSTITUIÇÃO DE BANCO DE DADOS PARA UMA HISTÓRIA SOCIAL LINGUÍSTICA DO BRASIL

As fontes remanescentes do passado são os únicos testemunhos de que dispõem os pesquisadores para o conhecimento da história de uma língua. Mattos e Silva (2002) já advertia que, para a reconstrução da sócio-história linguística do PB muitas histórias deveriam ser reconstruídas e que essas histórias dependiam, necessariamente, de fontes históricas de natureza diversa. Destaca-se que, em *Linguística Histórica*, “O uso de [fontes ou] *corpora* permite a realização de descrições

linguísticas de base empírica e promove, com isso, a discussão de questões teóricas solidamente fundamentadas.” (Nascimento, 2004, p. 1). Portanto, são os *corpora* essenciais em pesquisas desse campo de atuação.

Com relação ao PB popular, as fontes para a realização de pesquisas devem ser outras, especialmente porque o português característico das classes sociais mais baixas se fez na oralidade (Mattos e Silva, 2008), logo, materiais gráficos representativos dessa vertente são escassos, como destacou Santiago (2012).

Para a recuperação da história social e linguística dessa vertente, Mattos e Silva (2002, p. 24-25) entrevê “[...] o estudo vertical das variantes populares do português brasileiro, não só as urbanas, [...] mas, nas suas variedades rurais de todo o Brasil”. Como pauta de pesquisa para o PB popular no âmbito do PHPB, Mattos e Silva (2002, p. 461) propõe:

1. para a reconstrução do passado do português popular brasileiro, pesquisar no espaço brasileiro as variedades conviventes hoje sobretudo as dos não-escolarizados das diversificadas áreas rurais do Brasil;
2. reconstruir, detalhadamente e com a precisão possível, a sócio-história linguística das diversas áreas brasileiras cobertas pelo Projeto, considerando as línguas que aí estiveram em contacto, os movimentos demográficos, a ausência/presença da escolarização e, conseqüentemente da escrita, como elemento normatizador.

Lobo (2001, p. 109) traça o perfil do que seria um corpus ideal para o estudo do português popular brasileiro. A pesquisadora sugere que o corpus permita conhecer “as variedades populares – supostamente mais diversificadas e descendentes diretas das variedades do português como segunda língua”. No entanto, a busca por fontes escritas dos grupos socialmente marginalizados de épocas pretéritas não é uma tarefa das mais fáceis (Mattos e Silva, 2004), mesmo em se tratando do século XX, quando houve um impulso no processo de escolarização. Nas cidades mais afastadas e, sobretudo com uma densidade demográfica extremamente baixa, a implantação de escolas foi tardia e precária e não atendia a demanda existente, o que promoveu um processo de ensino irregular (Carneiro; Almeida, 2006).

Oliveira (2006) observa que os pesquisadores da área se esbarram em obstáculos na busca por documentos de grupos “subalternos”, especialmente de épocas mais recuadas, como a raridade, a dispersão arquivística e o mínimo quociente de durabilidade (Petrucci, 1999). Encontrar textos que representem a escrita de pessoas oriundas de contextos sociais desfavoráveis é fundamental para ajudar a contar a história do PB. Segundo Mattos e Silva (2008), é através da escrita remanescente do passado que se pode chegar à língua falada, uma vez que sem a fala não se escreve.

Dada a dificuldade de encontrar textos mais próximos do vernáculo, que reflitam uma escrita cotidiana, particularmente de pessoas menos escolarizadas, e de eras idas, localizá-los apresenta significativa relevância para os estudos em Linguística Histórica. Nesse sentido, a constituição de um banco de dados da vertente popular do PB, agenda à qual se filia este trabalho, é uma tarefa profícua e necessária no sentido de ajudar a contar a história da língua portuguesa, especialmente da região rural baiana.

3 O BANCO DE DADOS SOCIOLINGÜÍSTICOS DO NELP/UEFS

No Núcleo de Estudos de Língua Portuguesa (NELP) da UEFS, fundado, em 1998, pelas professoras doutoras Norma Lúcia Fernandes de Almeida e Zenaide de Oliveira Novais Carneiro, diversos projetos de constituição de *corpora* para o estudo da língua portuguesa vêm sendo desenvolvidos desde a sua concepção. Desde textos orais, como a constituição do banco de dados do projeto *A língua falada no Semi-Árido Baiano*, criado em 1993, com o auxílio do professor doutor Dante Lucchesi, a documentos escritos, como no projeto *Contribuições para a constituição de um banco de textos e de dados para o estudo da história do português no Brasil, do século XVII ao XX*, implantado em 1997, pela professora doutora Ilza Ribeiro e que, posteriormente foi filiado ao Programa para a História do Português (PROHPOR/UFBA), coordenado na época por Rosa Virgínia Mattos e Silva e ao *Projeto Para a História do Português Brasileiro* (PHPB), coordenado por Ataliba de Castilho (USP).

Atualmente, o NELP-UEFS vem desenvolvendo diversos projetos de pesquisas que buscam um efetivo conhecimento da realidade linguística brasileira, especialmente das regiões rurais e semiáridas baianas. Essas investigações, primordialmente, são fruto das inegáveis contribuições legadas por Rosa Virgínia Mattos e Silva aos pesquisadores que se propõem a investigar a história social e linguística do Brasil que, segundo a própria autora, “se moverá fundado na história social do Brasil, [...] aquele em que o historiador da língua estará mais próximo do historiador *tout court*” (Mattos e Silva, 2004, p. 59).

Dentre os projetos do NELP, destaca-se o projeto *Vozes do Sertão em Dados: história, povos e formação do português brasileiro*, do qual faz parte o banco *DOHS: Documentos Históricos do Sertão* e, a partir de 2010, também na versão eletrônica, o *CE-DOHS: Corpus Eletrônico de Documentos Históricos do Sertão*. Para a formação desse banco de dados, levaram-se em conta as proposições sugeridas por Rosa Virgínia Mattos e Silva e que foram extraídas de 10 textos de sua autoria, sobre a história social e linguística do PB, assim descritas por Lobo (2015, p. 70):

1. A história linguística do Brasil não se restringe à história da língua portuguesa no Brasil, nem à história do português brasileiro.
2. O português brasileiro emerge em contexto multilíngue: o contato linguístico é, pois, elemento constitutivo da sua formação¹.
3. Na cena linguística do Brasil colonial, destacam-se três atores principais: o português europeu, as línguas gerais indígenas e o português geral brasileiro.

¹ Sobre a constituição histórica do PB, Mattos e Silva (2004, p. 17) afirma que “a história das línguas passa necessariamente pela história demográfica de seus falantes”. Nesse sentido, considerando os dados de Mussa (1991), sobre o papel das línguas africanas na história do português do Brasil, nos quais apresenta uma síntese da demografia histórica do Brasil, a autora mostra que africanos, negros brasileiros, mulatos e índios integrados formavam uma maioria constante, correspondendo a 70% da população. Esse fato, segundo a autora, trouxe implicações linguísticas para a formação do PB, e acrescenta que “é certamente no entrecruzar-se de variantes localizadas menos ou mais interferidas por marcas indígenas e/ou africanas, de variantes mais gerais menos ou mais africanizadas ou menos ou mais aportuguesadas que se definem e emergem os traços característicos do português brasileiro, língua nacional”. Sendo assim, reafirmando a importância que teve o componente africano na formação do PB.

4. Africanos e afrodescendentes foram os principais difusores da língua portuguesa no Brasil e os principais formatadores do português brasileiro em sua variante social majoritária — o português popular brasileiro.

5. O passado sócio-histórico-linguístico do Brasil deverá ser interpretado para a compreensão do português brasileiro “heterogêneo e variável, plural e polarizado”.

Na descrição dos dados biográficos dos escreventes/informantes consideraram-se nome, sexo, data de nascimento, local de nascimento, etnia, escolaridade, grau de habilidade de escrita; em relação aos dados do documento/amostra, catalogaram-se número do documento, data de escrita, local de escrita, tipo de documento e fonte.

O CE-DOHS está organizado em duas fases: a primeira fase do projeto teve por objetivo compor um banco de textos de 1750 a 2000, representativos do período histórico do PB caracterizado pelo *multilinguismo localizado*. Esse banco permite estudar a história do português brasileiro socialmente prestigiado, semi-culto e socialmente estigmatizado, nesse contexto.

A segunda fase, iniciada em 2019, recua ainda mais no tempo, a uma época em que o *multilinguismo* no Brasil era *generalizado* (1500-1750) (Mattos e Silva, 2004; Lucchesi, 2017). Tem por objetivo estudar a gestação do PB. Essa fase enfrenta a raridade das fontes, pois são poucos os textos escritos por grupos nascidos no Brasil, sobretudo de índios e negros, etnias que não tiveram acesso à escola; o projeto, todavia, tem pequenos acervos desse período e bastante significativos.

O CE-DOHS, na era das humanidades digitais, busca contribuir com o PHPB, disponibilizando, em sua plataforma, textos em diferentes versões: manuscritos, impressos e amostras de fala. Como resultado da feliz parceria com o *Projeto Corpus Histórico do Português Tycho Brahe* (www.tycho.iel.unicamp.br), o CE-DHOS traz uma série de documentos, em sua maioria manuscritos, editados em linguagem modernizada, em uma versão XML (*eXtensible Markup Language*), com o uso da ferramenta *eDictor*, um editor de textos eletrônico, voltado especialmente para o trabalho filológico, que permite a geração automática de versões correspondentes a edições diplomáticas, semidiplomáticas e modernizadas (em *html*), e de versões com anotação morfossintática (em texto simples, e XML), desenvolvida por Sousa, Kepler e Faria (2007).

Como desdobramento do advento das novas tecnologias, no início do século XXI, surgiram novas ferramentas, hoje, fundamentais para os trabalhos no âmbito da Filologia, sem contudo, distanciar-se do rigor que a disciplina exige, essas ferramentas dispõem aos filólogos novas e múltiplas possibilidades de suporte, manipulação, processamento, armazenamento e apresentação do texto no meio digital, a exemplo, da ferramenta *eDictor*.

Como bem destacaram Gonçalves e Banza (2013, p. 4):

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

Os textos disponibilizados na plataforma do CE-DOHS foram produzidos por brasileiros de diferentes etnias, nascidos entre 1586 e 1986, e, também, por portugueses nascidos entre 1450 e 1850. São documentos encontrados em dezenas de arquivos públicos e privados, nacionais e internacionais; também gravações de fala em diversas regiões da Bahia. Atualmente, há 50 acervos disponíveis, que somam cerca de um milhão de palavras. A seguir, a relação de acervos compostos por cartas no âmbito do CE-DOHS:

Quadro 1 – Relação de acervos compostos por cartas no âmbito do CE-DOHS.

| Acervo | Descrição |
|---|---|
| Acervo Cartas para vários destinatários (1809-1904) | Traz 208 cartas (1-208, extraídas de Carneiro 2005, 2011), escritas entre 1809 e 1904, sendo 149 cartas escritas entre 1880 e 1899 e 38 cartas, entre 1900 e 1903, por 114 remetentes (111 homens e 3 mulheres), oriundos, em sua maioria, da classe alta e letrada. |
| Acervo Cartas para Cícero Dantas Martins, Barão de Jeremoabo (1880-1903) | Constitui-se de 190 cartas (extraídas de Carneiro, 2005). Escritas pela elite rural sertaneja (43 baianos; 42 homens e uma mulher). Essas cartas oferecem uma amostragem de textos escritos por brasileiros semicultos e populares, nascidos/radicados no interior, especificamente da Bahia, entre 1724-1880. |
| Acervo Cartas para Severino Vieira, Governador da Bahia (1901-1902) | Compõe-se de um maço de 102 cartas (209-310, extraídas de Carneiro, 2005), enviadas a Severino Vieira por 60 remetentes (57 homens e três mulheres), a maioria letrada e, sobretudo, cidadina. |
| Acervo Cartas para Dantas Jr. (1902-1962) | Traz 242 cartas enviadas para Dantas Jr. Ao todo são 113 remetentes, a maior parte é formada por Bacharéis em Direito com altos cargos no período republicano em que viveram. |
| Acervo Cartas em Sisal, Riachão do Jacuípe, Conceição do Coité e Ichu (1906-2000) | Constitui-se de 131 cartas pessoais escritas ao longo do século XX por 53 sertanejos oriundos da zona rural dos municípios de Riachão do Jacuípe, Conceição do Coité e Ichu, região sisaleira do semiárido baiano. A amostra é representativa da variedade popular do português brasileiro. |
| Acervo Cartas Particulares da Família Freire (1937-1942) | Pequeno conjunto de cartas, bilhetes e cartões manuscritos, trocados, entre Carlos Ribeiro Freire e Iracema Batista Chéquer Freire. São 17 correspondências, representativas de normas cultas do português brasileiro. |
| Acervo Cartas Particulares da Família Soledade (1948-1951) | Conjunto de 100 cartas manuscritas (apenas uma é datilografada), trocadas entre baianos cultos, Otto Soledade Júnior e Renée da Silva Barros Soledade. |
| Acervo Correspondências Amigas (1980-1993) | Constitui-se por 79 cartas, 25 cartões e 80 envelopes. São 38 remetentes, 31 do sexo feminino e sete do sexo masculino, a maior parte nascida/radicada no interior da Bahia, estudantes do Ensino Fundamental ou do Ensino Médio, falantes de um português semipopular ou semiculto. |
| Acervo da Família Oliveira (1962-1973) | São 23 cartas oriundas do Rio de Janeiro, sendo 22 apócrifas do recém-casado Arnaldo Andrade Dias, dirigidas a sua esposa Lourdinha [Maria de Lourdes Lima de Oliveira], na Bahia, e uma autógrafa de João Carvalho de Matos, “compadre” de Lourdinha. As cartas parecem ser de brasileiros falantes de um português semipopular. |
| Acervo Cartas Marienses (1935-1995) | Conjunto de 69 cartas, 17 cartões e cinco bilhetes, escritos entre 1935 e 1995, na região rural de Coração de Maria, Bahia, por 29 redatores, sendo 15 do sexo feminino e 14 do sexo masculino. Os remetentes, em sua maioria, possuem o nível primário de ensino, portanto da vertente socialmente estigmatizada do PB. |

O banco de dados do CE-DOHS resulta de pesquisas realizadas no âmbito do projeto e conta com a participação de professores pesquisadores, de bolsistas de

iniciação científica e de mestrandos e doutorandos que desenvolvem pesquisas nas áreas de ferramentas computacionais, humanidades digitais, prospecção, edição filológica, tratamento de *corpora* orais, tratamento metodológico de *corpora* escritos e metadados, e de *corpora* anotados. Conta também o projeto com a colaboração de pesquisadores de instituições parceiras e pesquisadores associados. A coleção *Cartas brasileiras* (Carneiro, 2011), coletânea de fontes para o estudo do português, em três volumes, resulta do trabalho que vem sendo desenvolvido no âmbito do CE-DOHS/NELP/UEFS.

4 FONTES PARA A HISTÓRIA DO PORTUGUÊS POPULAR BRASILEIRO

Lucchesi (2001) definiu a realidade sociolinguística brasileira como *bipolarizada*. O autor refere-se a “dois sistemas igualmente heterogêneos, daí a pluralidade e a polarização, que designará por norma culta e norma vernácula” (Mattos e Silva, 2002, p. 49). No polo das camadas médias e altas da sociedade, sobretudo nos três primeiros séculos de história, houve um comportamento conservador dos padrões linguísticos lusitanos por parte de uma reduzida elite colonial. Já no polo das camadas populares, ocorreram drásticas mudanças linguísticas, especialmente pelo intenso contato do português com as línguas indígenas e africanas, desencadeadas sobretudo através de um processo de *transmissão linguística irregular* (TLI) (Lucchesi, 2001, 2015; Lucchesi; Baxter, 2009). É possível pensar, inclusive, na emergência de variedades crioulistas no seio de comunidades africanas e afrodescendentes isoladas.

Esses acontecimentos sócio-históricos provocaram significativas consequências para a atual configuração do PB. Nesse contexto, mantiveram-se, de um lado dos polos, as normas cultas, propagadas pelo processo de escolarização e usadas, sobretudo, pelas classes prestigiadas; e, do outro lado, as normas populares, que foram disseminadas pelas falas dos grupos socialmente estigmatizados (Mattos e Silva, 2004).

Para Lucchesi (2015), as marcas mais drásticas dessa aquisição imperfeita, que resultaram nessa “babel” linguística, foram atenuadas a partir da terceira década do século XX, especialmente pela crescente industrialização e urbanização, pelo desenvolvimento dos sistemas de ensino e pelo desenvolvimento do sistema de comunicação em massa.

Sobre esta questão, Faraco (2008, p. 63) afirma:

A sociedade industrial moderna trouxe consigo uma série grande de efeitos, redesenhando a face do mundo contemporâneo [...]. Alteradas as condições objetivas de funcionamento da sociedade, alteram-se também as condições objetivas de funcionamento social da língua. A urbanização intensa, a expansão do sistema educacional, a formulação e difusão política do conceito moderno de cidadania e o desenvolvimento dos sistemas de comunicação em massa deram hegemonia e ampla difusão a certas variedades da língua, em particular às variedades tradicionalmente urbanas, que passaram a exercer poderosa força centrípeta sobre as demais variedades.

No entanto, mesmo com as alterações no funcionamento da língua, desencadeadas pelo amplo desenvolvimento social do país, ainda se mantém um “fosso” que separa as normas cultas e as normas populares. Lucchesi (2015), explica

que o desenvolvimento tardio, a superexploração do trabalho e a desigual concentração de renda impedem que haja um nivelamento linguístico na sociedade brasileira.

No CE-DOHS, destacam-se três acervos que disponibilizam fontes históricas remanescentes, constituídos por cartas, que representam o polo das normas populares, como as destacadas por Lucchesi (2001). Esses acervos, produzidos no século XX, são produtos gráficos de grupos sociais desfavorecidos, da região rural semiárida baiana, e trazem “as marcas que o contato linguístico produziu no passado” (Lucchesi, 2015, p. 151).

O corpus *Correspondências Amigas: o acervo de Valente (Bahia)* é um acervo constituído por 79 cartas, 25 cartões e 80 envelopes. Foi escrito por 38 remetentes, 31 do sexo feminino e sete do sexo masculino. Os remetentes, a maior parte nascida/radicada no interior da Bahia, jovens, com idade entre 20 e 30 anos, estudantes do Ensino Fundamental ou do Ensino Médio. O material permite a análise de fatos linguísticos diversos e representa o português escrito na segunda metade do século XX, por remetentes semicultos, semi-populares e populares. O material foi organizado por Oliveira et al. (2011), e publicado em CD-ROM, pela editora da UEFS.

O corpus *Cartas em Sisal* é constituído por 131 cartas pessoais escritas ao longo do século XX, por 53 sertanejos oriundos da zona rural dos municípios de Riachão do Jacuípe, Conceição do Coité e Ichu, localizados na região sisaleira do semiárido baiano. Os remetentes são lavradores e possuem pouca escolarização. Um conjunto de características presentes nas cartas fornece algumas pistas para perceber que os seus autores são indivíduos pouco familiarizados com a língua escrita, por isso a amostra revela-se representativa da variedade popular do PB. Esse acervo, organizado por Santiago (2011), integra o volume três da coletânea *Cartas brasileiras (1809–2000): coletânea de fontes para o estudo do português*, publicado pela editora da UEFS.

O corpus *Cartas Marienses (séc. XX)* é constituído por um conjunto de 69 cartas, 17 cartões e cinco bilhetes, escritos, entre 1935 e 1995, por 29 redatores oriundos da região rural de Coração de Maria, trabalhadores rurais, sendo 15 do sexo feminino e 14 do sexo masculino. Os remetentes, em sua maioria, possuem o nível primário de ensino, portanto pouco familiarizados com as normas gramaticais disseminadas pelo processo de escolarização. A documentação inédita, do século XX, permite o estudo e análise de diversos fenômenos linguísticos, sobretudo das vertentes socialmente estigmatizadas do PB.

Esses acervos, representativos das vertentes populares do PB, configuram uma importante e rica documentação, fontes de material empírico que podem trazer esclarecimentos sobre a penetração e difusão da língua escrita no interior baiano.

5 CARTAS MARIENSES: CONTEXTUALIZAÇÃO SÓCIO-HISTÓRICA

Segundo Mattos e Silva (2004, p. 110), no âmbito da Linguística Histórica, para um melhor controle metodológico de *corpora* histórico-diacrônicos, faz-se necessário contextualizar a amostra com dados da dimensão interna e externa da escrita. Para tanto, o pesquisador terá que controlar as seguintes proposições: *O que escreveu? Quando escreveu? Onde escreveu? Para quem escreveu? Afinal de contas, quem escreveu?* Essas proposições estão consoantes à proposta de Petrucci (2003, p. 7-8), para quem, no âmbito da Cultura Escrita, para qualquer tempo histórico, deve ser respondido um conjunto mínimo de questões, a saber:

- i) *¿Qué?* En Qué consiste el texto escrito, qué hace falta transferir al código gráfico habitual para nosotros, mediante la doble operación de lectura y transcripción;
- ii) *¿Cuándo?* Época en que el texto en sí fue escrito el testimonio que estamos estudiando;
- iii) *¿Dónde?* Zona o lugar en que se llevó a cabo la obra de transcripción;
- iv) *¿Cómo?* Con qué técnicas, con qué instrumentos, sobre qué materiales, según qué modelos fue escrito ese texto;
- v) *¿Quién lo realizó?* A qué ambiente sociocultural pertenecía el ejecutor y cuál era en tiempo y ambiente la difusión social de la escritura.
- vi) *¿Para qué fue escrito ese texto?*Cuál era la finalidad específica de ese testimonio en particular y, además, cuál podía ser en su época y en su lugar de producción la finalidad ideológica y social de la escritura.

O corpus foi contextualizado da seguinte forma:

a) *¿Qué?*

O corpus *Cartas Marienses – BA (séc. XX)* é uma documentação epistolar, de caráter pessoal, inédita, constituída por um conjunto de 67 cartas, – (a maioria autógrafas), 17 cartões e cinco bilhetes. O material integra o CE-DOHS do projeto *Vozes do Sertão em dados: história, povos e formação do português brasileiro*, um dos projetos do NELP-UEFS.

A referida documentação encontra-se em bom estado de conservação. Alguns documentos apresentam pequenas manchas, corrosões, marcas de dobras, partes amassadas e bordas levemente fragmentadas. Outros apresentam rabiscos, desenhos, adições e anotações; feitas, supostamente, por outra mão. Há também desgastes, por conta do manuseio, da umidade dos locais onde foram arquivados, sobretudo pelo tempo de uso.

As correspondências que compõem esse corpus, segundo Barbosa (1999, p. 147)², podem ser consideradas de circulação privada. Segundo o autor, em se tratando de textos escritos no passado, essa tipologia textual é o desejo de consumo de todo investigador em Linguística Histórica, pois é “[...] o que mais se aproxima de uma escrita cotidiana, aquela que tem maior chance de alcançar o limite possível de transparência na escrita de dados da oralidade, ocorrências de formas novas em difusão na sociedade”. Desse modo, textos valiosos para a compreensão sociolinguística da região.

b) *Quando?*

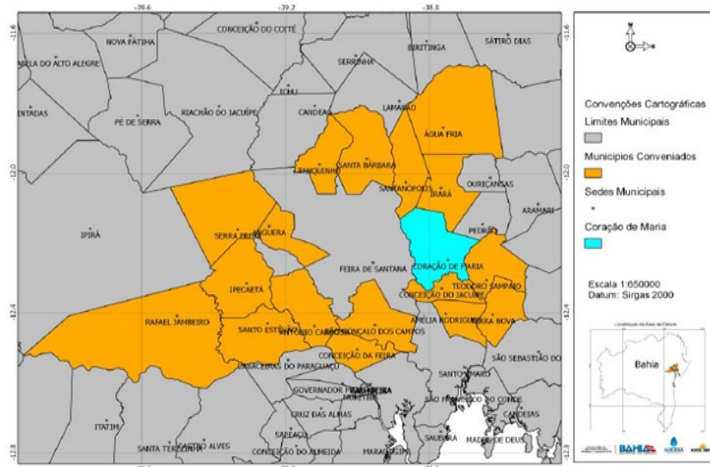
As correspondências foram escritas entre 1935 e 1995. Nas décadas de 30, 40, 50 e 60, há um número pequeno de cartas; esse fato pode ser explicado pela escassez de escolas na região, circunstância que impossibilitou o acesso da maioria da população à cultura letrada.

Há, no acervo, 52 cartas não datadas; no entanto, é possível que se faça uma aproximação da data de escrita, sobretudo pelo conteúdo das correspondências, que trazem evidências que contextualizam aspectos da época.

² Barbosa (1999) classifica os documentos do período colonial em três critérios ou macro categoriais, a saber: documentos da administração pública, privada e particular. Os documentos de circulação oficial incluem os documentos da administração pública e os documentos de circulação privada incluem os documentos da administração privada e os particulares.

c) *Onde?*

As correspondências foram produzidas por redatores³, oriundos do sertão baiano, especificamente moradores da região rural de Coração de Maria. O município localiza-se na Zona Fisiográfica de Feira de Santana, está parcialmente incluído no Polígono das Secas e integra o Polígono Portal do Sertão. O povoado surgiu em meados do século XIX, mas, só em 1944, foi emancipado. A sociedade Mariense é formada por descendentes de italianos e portugueses e, embora não conste nos dados históricos do município, é certo que outros contingentes populacionais, a exemplo de indígenas, escravos e seus descendentes, povoaram a região. A seguir, o mapa da região do município:



Fonte: Plano Municipal de Saneamento Básico de Coração de Maria - Ba - 2016.

Figura 1 - Mapa com a localização do município.

As principais atividades econômicas desenvolvidas pelos habitantes do município estão relacionadas à agricultura e pecuária, como a plantação de pequenas lavouras, como a mandioca, o amendoim, o milho, o abacaxi, a criação de pequenos e grandes animais de corte.

O município possui uma área de 372,315 km² de extensão, possui uma população de 22.041 habitantes e densidade demográfica de 64,34 hab./km². Desse total de habitantes, 10.841 são do sexo masculino e 11.560, do sexo feminino. Desses, 58% se concentram na zona rural, e os outros 42%, na zona urbana (IBGE).

Sobre o perfil sociolinguístico do município, há indícios de que houve, desde o início de sua formação, uma polarização sociolinguística. De um lado, a língua culta usada pela classe social privilegiada e que tinha acesso ao ensino formal e, do outro, os falares populares disseminados pela gente pobre que habitava a região.

Das 89 correspondências, em 30 há indicação de local de escrita, das quais 21 foram escritas na zona rural do município; as outras foram escritas no povoado de Laranjeira, Bento Simões, município de Irará, e as nove cartas restantes foram escritas em Salvador, por moradores do município que estavam residindo na capital baiana na

³ Utiliza-se o termo *redatores* para caracterizar os indivíduos que escreveram/produziram/redigiram as correspondências, ou seja, o termo foi empregado no sentido de autor da redação/do texto escrito, ou seja, redator.

data de escrita. Como é amplamente conhecido, a busca por melhores condições de vida impulsionou o êxodo rural em direção aos grandes centros urbanos, sobretudo a partir da década de 1930, com o desenvolvimento industrial e urbano do país (Lucchesi, 2015).

Com relação às correspondências que não tiveram seus locais de escrita especificados, infere-se, pela relação com as demais correspondências, pelo conteúdo expresso e também pelo local de residência dos redatores, que foram escritas na zona rural do município, no centro da cidade e em Salvador.

d) Como?

As correspondências foram escritas em papel comum, especificamente, em folhas de cadernos, outras em papel ofício, as demais em papel cartão. Algumas cartas foram escritas em papéis reutilizados. O instrumento utilizado para a escrita foi caneta esferográfica nas cores, azul, preta, vermelha e lilás; caneta hidrográfica lilás e lápis. As tipologias textuais empregadas foram: a carta, o cartão e o bilhete. Em algumas correspondências há poemas.

e) Quem?

Para um melhor controle metodológico da amostra, foi necessária a construção do perfil biográfico dos remetentes. Essas informações foram obtidas por meio de entrevistas com conhecidos, parentes, amigos, geralmente as pessoas para quem as correspondências foram destinadas. Nessa entrevista, o objetivo foi fazer um levantamento de informações que pudessem esclarecer quem foram esses remetentes. Os dados levantados foram catalogados em fichas, segundo o modelo utilizado por Carneiro (2005), como se observa no Quadro 2, a seguir:

FLP 24(1)

Quadro 2 - Dados pessoais dos remetentes.

| |
|---|
| Nome (conforme a carta): Egidio Mendes de Oliveira |
| Nome completo: Egídio Mendes de Oliveira |
| Filiação: João Mendes de Oliveira / Loriania Felícia de Jesus |
| Avós paternos: Ignorados |
| Maternos: José da Silva Rios/ Maria Macrina de Jesus |
| Naturalidade: Coração de Maria Nacionalidade: Brasileiro |
| Data de nascimento: 1938 Data de falecimento: Há pelo menos 10 anos (2009) |
| Idade do remetente (quando da escrita da carta): Informação desconhecida |
| Estado civil: Solteiro quando da escrita da carta |
| Instituição de ensino: Nível primário de escolarização |
| Profissão por formação: Informação desconhecida |
| Principais atividades: Informação desconhecida |
| Observações: Nasceu na Fazenda Santa Rosa, em coração de Maria. Lá aprendeu as primeiras letras juntamente com seus irmãos, em ambientes extraescolares. Deslocou-se para Salvador quando tinha pelo menos 30 anos de idade, lá permaneceu alguns anos, onde cursou o ensino primário de escolarização. Posteriormente, foi morar em São Paulo até o ano de sua morte. |
| Fonte: Depoimentos concedidos por sua sobrinha Marilene Mendes de Oliveira |

Fonte: Carneiro (2005).

O preenchimento dessa ficha teve o objetivo de fazer um levantamento sócio-histórico do remetente. Nela é possível identificar o remetente, a origem, a naturalidade e nacionalidade, o período histórico em que nasceu e viveu, a faixa etária do remetente

no período de escrita da correspondência, o nível de escolaridade e formação, bem como as atividades profissionais desempenhadas.

São 29 remetentes, 15 mulheres e 14 homens, brasileiros, nascidos entre 1880 e 1970, trabalhadores rurais, oriundos de Coração de Maria e região, que possuíam, em sua maioria, o nível primário de escolarização. A seguir, o Quadro 3 traz informações que caracterizam os remetentes, no que diz respeito à identificação do remetente, o nível de escolaridade que cada um possui, a etnia a qual faz parte, a data de nascimento, o local de nascimento, a data de escrita da carta e a quantidade de cartas escritas por cada remetente.

Quadro 3 – Caracterização dos remetentes. (*) Por inferência⁴.

| Remetente | Nível de escolaridade | Etnia | Data de Nascimento | Local de Nascimento | Data de escrita da carta | Quantidade de cartas |
|-------------------------------|--------------------------------|--------------|--------------------|---------------------|--------------------------|----------------------|
| Abdias Dias de Leão | Formação primária | Negra | 1918 | Coração de Maria | 15/02/1941 | 1 |
| Ana de Jesus Nogueira Martins | 4ª série do ensino primário | Branca | 1939 | Coração de Maria | 1992* | 1 |
| Antônio Muricy de Oliveira | Nível primário | Branca | 15/04/1940 | Coração de Maria | 1987* | 4 |
| Antônio Onofre dos Santos | Sabia ler e escrever | Negra | 1909 | Coração de Maria | 6 /03/ 1940 | 1 |
| Doraci Cerqueira Martins | 4ª série do ensino primário | Branca | 23/ 07/1932 | Coração de Maria | 1994 | 3 |
| Edinice C. Silva | Ensino Médio | Desconhecida | 1958* | Coração de Maria | 20/12/1980 | 1 |
| Edvaldo | Formação primária* | Desconhecida | 1970* | Coração de Maria | 1993 | 1 |
| Egídio Mendes de Oliveira | Formação primária | Branca | 1938 | Coração de Maria | 1964 | 1 |
| Feliciano Pereira | Analfabeto | Desconhecida | 1910* | Bento Simões | 1935 | 1 |
| Helena Almeida | Formação primária | Branca | 1946* | Coração de Maria | 1995 | 1 |
| Isabel dos Santos Lima | Formação primária | Branca | 1946 | Coração de Maria | 1988 | 1 |
| José Carlos Rodrigues Brandão | 7ª série do ensino fundamental | Branca | 1957 | Coração de Maria | 1980* | 2 |
| José Figueiredo de Miranda | Formação primária | Branca | 1890* | Coração de Maria | 1941 | 1 |
| José Mendes de Oliveira | Formação primária | Branca | 1935 | Coração de Maria | 1954/1959 | 2 |
| Joselito Pacheco da Silva | Ensino Médio completo | Branca | 1946 | Coração de Maria | Entre 1977 e 1980 | 8 |

⁴ O asterisco foi usado para inferir informações, as quais não foram possíveis fazer o levantamento.

| | | | | | | |
|----------------------------------|-----------------------------|--------------|-------|----------------------|--------------------|----|
| Madalena Nogueira Martins | 4ª série do primário | Branca | 1951 | Coração de Maria | 1993 | 3 |
| Manoel Ermenegildo Pereira Silva | Ensino primário | Desconhecida | 1939 | Coração de Maria | 1968 | 1 |
| Manoel Estevam Pacheco | Analfabeto | Negra | 1880* | Bento Simões – Irará | 1942 | 1 |
| Margarete Nogueira Martins | 8ª série ensino fundamental | Branca | 1972 | Coração de Maria | 1993 | 2 |
| Maria Natividade Pacheco Silva | Formação primária | Negra | 1916 | Bento Simões – Irará | 1969* | 1 |
| Maria Jose Pacheco da Silva | Formação primária | Negra | 1915 | Bento Simões – Irará | Entre 1984* e 1995 | 39 |
| Maria Cardoso Leal | Aprendeu a ler e escrever | Branca | 1938 | Coração de Maria | 1960 | 1 |
| Maria Santos de Jesus | Formação primária | Branca | 1946 | Coração de Maria | 1973 | 1 |
| Maria Zélia Nogueira Martins | 4ª série do primário | Branca | 1954 | Coração de Maria | Entre 1984 e 1990* | 6 |
| Mocinho | Ensino primário* | Desconhecida | 1920* | Coração de Maria* | 1964 | 1 |
| Nilza | Ensino Médio | Desconhecida | 1960 | Coração de Maria | 1984 | 1 |
| Rizonete Nogueira Martins | 4ª série do primário | Branca | 1970 | Coração de Maria | 1992* | 1 |
| Zeferino | Ensino primário* | Desconhecida | 1925* | Coração de Maria | 1954 | 1 |

f) *Para quê?*

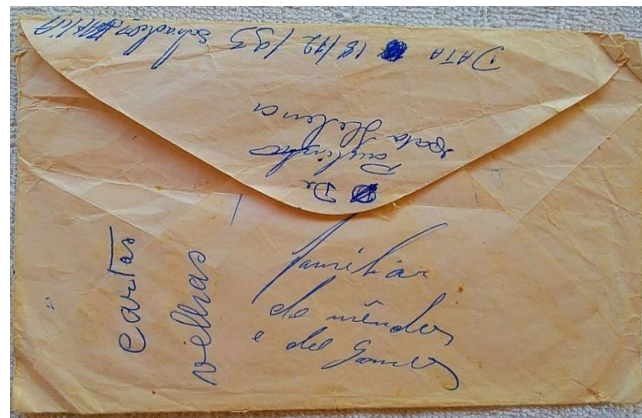
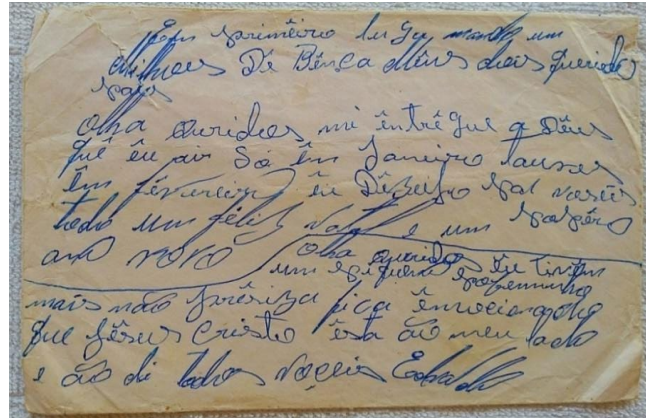
São correspondências trocadas entre familiares, amigos e conhecidos, cuja finalidade era expressar saudades, dar e receber notícias familiares, fazer pedidos, desejar felicitações de aniversário e boas festas, fazer convites, entre outros. Foram enviadas, em sua maioria, por portador, visto que muitos dos destinatários moravam na mesma comunidade. Apenas algumas correspondências foram enviadas via serviços postais.

6 AS EDIÇÕES FAC-SIMILAR E SEMIDIPLOMÁTICA DAS CARTAS MARIENSES

Segundo Borges e Souza (2012, p. 33), a edição fac-similar “é a reprodução fotográfica, transferindo-se a imagem do documento para o meio digital, ou convertendo a imagem, ou sinal dialógico para o código digital, realizando-se a digitalização dos textos”. A edição fac-similar foi o primeiro passo dado após a localização da documentação. Segundo Lose (2017), essa edição é considerada de baixo nível de mediações. No processo de edição, cabe ao editor organizá-la e elaborar paratextos, informações complementares e análises.

A edição fac-similar, respeitando técnicas e condições adequadas de execução, preserva a autenticidade do manuscrito e facilita o processo de edição semidiplomática,

uma vez que, no meio digital, é possível que o documento seja aproximado para uma melhor visualização. As fotografias dos manuscritos foram feitas em maio de 2018. Os documentos foram colocados abertos sobre uma mesa na qual foram captadas as imagens. As Figuras 2 e 3, respectivamente, trazem o fac-símile da carta 89 (E), escrita em um envelope, por Edvaldo, na década de 1990.



Fonte: CE-DOHS.

Figuras 2 e 3 - Fac-símile da carta 89 [recto e verso].

A edição semidiplomática, conservadora, preserva as características linguísticas do manuscrito, tão fundamentais para o trabalho em Linguística Histórica. Essa edição apresenta um nível moderado de mediações (Lose, 2017), por isso “procura deixar o texto o mais fiel possível, [...] as interferências são previamente estabelecidas, as quais permitem que as características linguísticas e ortográficas sejam mantidas” (Queiroz, 2012, p. 16).

Sobre a edição filológica, Mattos e Silva (2008, p. 15) explica que “[...] não se pode nem se deve utilizar qualquer edição de texto do passado para a análise histórico-diacrônica: a edição tem de ter sido feita com rigor filológico e com o objetivo claro de servir a estudos linguísticos”. Em consonância com a autora, Maia (2012, p. 536, 537), ao relacionar Filologia e Linguística Histórica, afirma que é íntima a relação entre as duas ciências e que para a investigação em perspectiva diacrônica, “é absolutamente necessária uma consistente infraestrutura filológica”.

A edição semidiplomática, conservadora, foi feita a partir dos fac-símiles das correspondências. Os critérios foram elaborados conforme as normas de transcrição para documentos manuscritos do PHPB (Mattos e Silva, 2001), relacionadas a seguir:

1. Cabeçalho:

- a) Identificação do arquivo.
- b) Indicação do número de fólios, das dimensões do suporte, do estado de conservação e empaginação.
- c) Indicação de estampilhas, marca d'água, timbre, ornamentos.
- d) Anotações do arquivo feitas nos documentos

2. Marcas supragráficas serão indicadas em nota de rodapé de acordo com as especificações a seguir:

- a) [.] Parte do documento não transcrita com indicação em nota do tipo de problema encontrado: rasuras, borrões, dano no suporte etc.
- b) [?] Leitura difícil de vocábulo.
- c) [??] Trecho de difícil leitura.

3. As abreviaturas, alfabéticas ou não, serão desenvolvidas, marcando-se, em itálico, as letras omitidas na abreviatura, obedecendo aos seguintes critérios:

- a) Respeitar, sempre que possível, a grafia do manuscrito, ainda que manifeste idiossincrasias ortográficas do redator.
- b) No caso de variação no próprio manuscrito, a opção será para a forma atual ou mais próxima da atual, como no caso de ocorrências do tipo “Coraçam” e “Coração”, que levam a abreviatura “C” a ser transcrita “Coração”.

4. Não será estabelecida fronteira de palavras que venham escritas juntas, nem se introduzirá o hífen ou apóstrofo onde não houver. Exemplo: “Antonio agonia aplantou amendoin com a terra seca | agora está de cabêça quente eu não estou me | encomodando que Deus sabe o que faz”.

5. A pontuação e a acentuação originais serão rigorosamente mantidas. Exemplo: “Meu querido irmão zezinho eu vô ao fim desta | e para dar as minha noticia i ao mesmo tempo | obiter as suas meu irmão eu fui bem de viagi grassa | ao nosso bom deus meu irmão eu vô bem isto com saudi | e você zezinho vai com saudi fasso votos que você [...]”.

6. Será respeitado o emprego de maiúsculas e minúsculas como se apresentam no original. No caso de algumas variações físicas dos sinais gráficos resultarem de fatores cursivos, não serão consideradas relevantes. Assim, a comparação do traçado da mesma letra deve propiciar a melhor solução. Exemplo: “Vamos ama | A Vida Por- que na | Vida Deus no ama | So Deus que nois | Da Foça e Paz para | nois Vencer todas | as Dificolidade que[...]”.

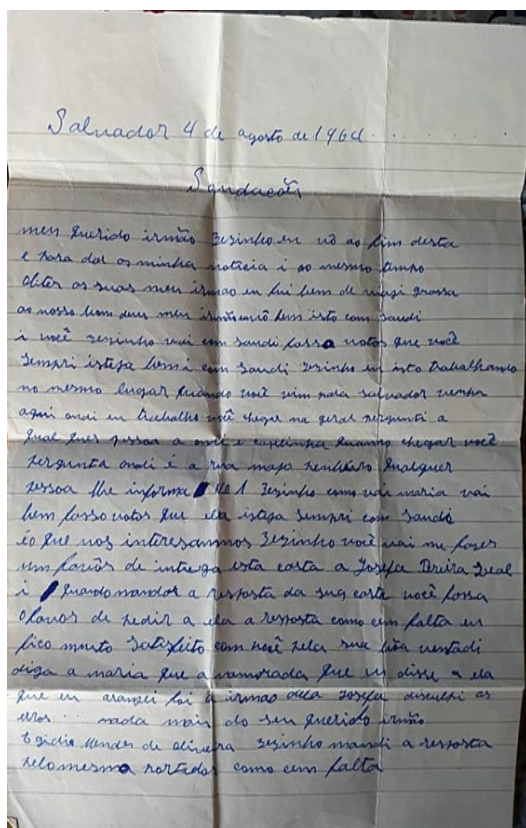
7. Supressões feitas pelo redator no original serão tachadas. Exemplo: “Rute Modobi ficou bom logo no | outro dia, obrigada. | ~~Hansil está ficando pouco~~”.

8. A divisão das linhas do documento original será preservada. A indicação será feita na edição pela marca de uma barra vertical | entre as linhas. A mudança de fólio receberá a marcação com o respectivo número na sequência, 1v., 2r., 2v., 3r., etc.

9. Inserções do redator, se na entrelinha, entram na edição entre os sinais <↑>, se na entrelinha superior; <↓>, se na entrelinha inferior; se nas margens superior, laterais ou inferior, entram na edição entre os sinais < >, na localização indicada. Exemplo: “<Quando acharem um caderno só para mim>”.

10. A assinatura virá sublinhada. Exemplo: Abedias Dias di Leão. Serão indicados em nota, qualquer sinal idiossincrático ou quando a assinatura não corresponder à grafia da carta.

A Figura 4 traz o fac-símile e a edição semidiplomática da carta 58 do acervo, escrita por Egidio Mendes de Oliveira, em 1964:



Carta 58

AJMO. Documento contendo um fólio. Escrito com tinta azul, em papel almaço com pautas. O fólio apresenta marcas de dobras

Salvador 4 de agosto de 1964.....|

Saudações|

meu querido irmão zezinho eu vô ao fim desta | e para dar as minha noticia i ao mesmo tempo | obiter as suas meu irmao eu fui bem de viaji grassa | ao nosso bom deus meu irmão eu vô bem isto com saudi | e você zezinho vai com saudi fasso⁵ votos que você | sempri isteja bem i com saudi zezinho eu isto trabalhando | no mesmo lugar quando você vim para salvador venha | aqui ondi eu trabalho você chega na geral pergunti a | qual quer pessoa a ondi e capelinha quando chegar você | pergunta ondi é a rua majo pinheiro⁶ qualquer | pessoa lhe informa [.]⁷ Número 1 zezinho como vai maria vai | bem fasso votos que ela esteja sempri com saudi⁸ | é o que nos interesamos zezinho você vai me fazer | um favôr de intrega esta carta a Josefa Pereira Leal | i [.]⁹ quando mandar a resposta da sua carta você fassa | o favor de pedir a ela a resposta como cem falta eu | fico muito subjeito com esse peca sua boa vontade diga a maria que a amparada que eu disse a ela que eu arranji foi a irmao dela sorija desculpi os mas ... nada mais do seu querido irmão | Egidio Mendes de oliveira zezinho mandei a resposta pelo mesma¹⁰ portador como cem falta

FLP 24(1)

Fonte: CE-DOHS.

Figura 4 – Edição semidiplomática com fac-símile da carta 58 (EMO).

6.1 Edição modernizada

A partir da edição semidiplomática, o CE-DOHS desenvolve a edição modernizada, utilizando, como ferramenta, o *eDictor*, um editor de textos, desenvolvido por Sousa e Kepler (2007) e Sousa, Kepler e Faria (2013), voltado para o tratamento filológico e a análise linguística automática. A ferramenta combina um editor de XML e um etiquetador morfossintático, que permite a geração automática de edições nas versões diplomáticas, semidiplomáticas e modernizadas, e de versões com anotação morfossintática. No processo de edição, diferentes camadas de edição são aplicadas ao documento, o que permite a recuperação das palavras grafadas originalmente.

⁵ Rasurado
⁶ Rasurado
⁷ Rasurado
⁸ Rasurado
⁹ Rasurado
¹⁰ Rasurado

A versão modernizada pode padronizar os elementos convencionais da escrita, como a grafia, a acentuação, o desenvolvimento de abreviaturas. Porém, todas essas alterações ficam visíveis, o que permite que o leitor possa acessar o documento original, tendo controle das intervenções feitas no texto. Na edição digital, mantêm-se as mudanças de parágrafo, a orientação da escrita, as mudanças de linha, as correções do autor, entre outros.

As Figuras 5, 6, e 7 trazem, respectivamente, a versão semidiplomática, modernizada e itens modernizados da carta 58, gerada pelo *eDictor*:



Fonte: CE-DOHS.

Figura 5 - Edição semidiplomática, com fac-símile, gerada pelo *eDictor*.

FLP 24(1)



Fonte: CE-DOHS.

Figura 6 - Edição modernizada final, gerada pelo *eDictor*.

A Figura 7 traz uma relação de dados que foram modernizados na carta 58 do acervo. Observa-se a padronização da grafia nas palavras: *graça* por *grassa*, *obter* por *obiter*, *onde* por *ondi*, *Maria* por *maria*, *desculpe* por *desculpi*, *sem* por *cem*, *Major* por *majo*,

entre outras; e também, no que diz respeito à acentuação: *é* por *e*, *notícia* por *noticia*, *favor* por *favôr*.

Carta pessoal 58-EMO-4-09-1964

- [1] Autor: Egídio Mendes de Oliveira
- [2] Destinatário: José Mendes de Oliveira (Zezinho)
- [3] Data: 04/08/1964
- [4] Lista dos itens editados (por ordem alfabética)
- [5] ver ficha catalográfica e outras versões disponíveis

| Item Original | Edições |
|---------------|------------------------|
| [] | ras: [j], |
| [] | ras: [j], |
| bôa | pad: boa, |
| cem | pad: sem, |
| deus | pad: Deus, |
| disculpi | pad: desculpe, |
| e | pad: é, |
| fassa | pad: faça, |
| fasso | pad: faço, ras: fasso, |
| fav-ôr | pad: favor, |
| grasia | pad: graças, |
| i | pad: e, |
| i | pad: e, |
| i | pad: e, |
| intrega | pad: entregar, |
| irmas | pad: irmãs, |
| isteja | pad: esteja, |
| isto | pad: estou, |
| majo | pad: Major, |
| mandi | pad: mande, |
| mama | pad: Mãe, |
| mama | pad: Mãe, |
| mesma | ras: mesma, |
| meu | pad: Meu, |
| meu | pad: Meu, |
| noticia | pad: notícia, |
| obiter | pad: obter, |
| ondi | pad: onde, |
| ondi | pad: onde, |
| ondi | pad: onde, |
| pergunti | pad: pergunte, |

Fonte: CE-DOHS.

Figura 7- Lista de dados modernizados.

O corpus *Cartas Marienses* está disponível na Plataforma *Corpus Eletrônico de Documentos Históricos do Sertão* (CE-DOHS) (www.uefs.br/cedohs/), em diferentes versões de apresentações, acompanhadas dos fac-símiles dos documentos: semidiplomática e modernizada¹¹.

É possível acessar informações completas sobre o acervo no endereço: www.uefs.br/cedohs/cartasmarienses/cartas/¹². Nesse espaço, estão disponibilizadas a documentação completa, em diferentes versões de edição, as informações sobre o perfil biográfico dos remetentes, o índice analítico das cartas e as produções feitas a partir do acervo.

¹¹A edição modernizada foi realizada por Priscilla Tuy Batista, pesquisadora integrante do CE-DOHS.

¹²O site foi desenvolvido pelo engenheiro da computação Igor Leal Souza, integrante da equipe do CE-DOHS.

Cartas Marienses Início Cartas Produções Buscar

Sobre as cartas

Este espaço foi criado com o objetivo de disponibilizar o Acervo denominado "Cartas Marienses". A amostra é constituída por um conjunto de 70 cartas, 17 cartões e 5 bilhetes, escritos, entre 1935 e 1995, na região de Coração de Maria- Bahia, por 30 redatores, sendo 16 do sexo feminino e 14 do sexo masculino. Os remetentes, em sua maioria, possui o nível primário de ensino, portanto pouco familiarizados com as normas gramaticais disseminadas pelo processo de escolarização. São cartas, principalmente, de caráter afetivo, enviadas com o objetivo de expressar saudades, obter e dar notícias, fazer pedidos, enviadas por mães, filhos, irmãos, primos, amigos, compadres, vizinhos, conhecidos, entre outros. A documentação inédita, do século XX, permitirá o estudo e análise de diversos fenômenos linguísticos, sobretudo das vertentes socialmente estigmatizadas do português brasileiro. Detruçar-se sobre estes escritos é resgatar elementos fundamentais para o entendimento de uma história linguística e social do português brasileiro do passado, especificamente desta região. Além do acervo, são disponibilizadas: as informações sobre o perfil biográfico dos remetentes; a autobiografia da remetente Maria José Pacheco e as produções feitas a partir do acervo. O corpus Cartas Marienses – BA (séc. XX) faz parte do acervo do banco Corpus Eletrônico de Documentos Históricos do Sertão – CE-DOHS (www.uefs.br/cedohs), do Projeto Vozes do Sertão em Dados: história, povos e formação do Português Brasileiro (CNPq. 401433/2009-9), coordenado pela Profa. Dra. Zenaide de Oliveira Novais Carneiro e pela Profa. Dra. Mariana Fagundes de Oliveira Lacerda, um dos projetos do Núcleo de Estudos em Língua Portuguesa (NELP), da Universidade Estadual de Feira de Santana (UEFS), na Bahia.

Acervo

São 70 cartas, 17 cartões, 5 bilhetes, de caráter pessoal, editados na versão semidiplomática, com fac-símile, de acordo com critérios estabelecidos pelo projeto Para a História do Português Brasileiro (PHPB). A documentação encontra-se também editadas na versão modernizada, em linguagem XML, com o uso da ferramenta eDíctor. Para ter acesso ao acervo, [clique aqui](#).

Redatores

São 30 remetentes, 16 mulheres e 14 homens nascidos no município de Coração de Maria e região. São trabalhadores rurais e donas de casas que vivem da plantação de pequenas lavouras como: o abacaxi, a mandioca, o milho, fumo, amendoim, o coco, a banana, entre outras. Alguns destes remetentes, quando da escrita das correspondências, se encontravam em Salvador, principal centro migratório das pessoas da região, que buscavam melhores condições de vida.

Fonte: www.uefs.br/cedohs/cartasmarienses/cartas/.

Figura 8 - Endereço eletrônico do acervo *Cartas Marienses*.

7 PALAVRAS FINAIS

A reconstrução da história social e linguística do PB, como afirma Mattos e Silva (2004, p. 61), “será trabalho para muitas mãos, durante muito tempo”. A pedra angular para a realização de tal propósito são as fontes escritas remanescentes do passado. Por isso, constituir bancos de dados sociolinguísticos, com base nos quais se possa contar essa história, é fundamental; e são de muita relevância bancos que tragam documentação que represente a escrita dos grupos que, historicamente, foram desfavorecidos.

O CE-DOHS, pioneiro do Nordeste, pertencente ao NELP, maior núcleo de estudos de Língua portuguesa, em perspectiva histórica, da Bahia, disponibiliza, em sua plataforma, diversos *corpora*, que foram tratados com rigor filológico, unindo a Antiga e a Nova Filologia, na era das humanidades digitais (Gonçalves; Banza, 2013). O acervo *Cartas Marienses*, cujas edições semidiplomática e modernizada estão aqui apresentadas, é um desses *corpora* que atendem à agenda de estudos sobre as normas socialmente estigmatizadas do PB.

REFERÊNCIAS

Borges R, Souza AS. Filologia e edição de texto. In: Borges R, et al., organizadores. Edição de texto e crítica filológica. Salvador: Quarteto editora; 2012. p. 15-59.

Carneiro ZON. Cartas brasileiras (1808-1904): um estudo lingüístico-filológico [tese]. Campinas: Instituto de estudos da linguagem, Universidade Estadual de Campinas; 2005.

Carneiro ZON, organizadora. Cartas brasileiras (1809-2000): coletânea de fontes para o estudo do português. Vol. 1 (1809-1904). Feira de Santana: UEFS editora; 2011.

Carneiro ZON, Almeida NLF. A criação de escolas a partir de critérios demográficos na Bahia do século XIX: uma viagem ao interior. In: Lobo T, Ribeiro I, Carneiro, Z, organizadores. Para a história do português brasileiro: novos dados, novas análises. Vol. 2. Salvador: EDUFBA; 2006. p. 649-673.

Cartas Marienses. [citado 20 jul. 20]. Disponível em: <http://www5.uefs.br/cedohs/cartasmarienses>.

- Corpus CE-DOHS: corpus eletrônico de documentos históricos do sertão. [citado 26 abr. 2020]. Disponível em: www.uefs.br/cedohs.
- Faraco CA. Norma culta brasileira: desatando alguns nós. São Paulo: Parábola editorial; 2008.
- Gonçalves MF, Banza AP. Fontes de metalinguísticas para a história do português clássico. In: Gonçalves MF, Banza AP. Patrimônio textual e humanidades digitais: da antiga à nova filologia. Évora: CIDEHUS; 2013. p. 73-112.
- Instituto Brasileiro de Geografia e Estatística (IBGE). Cidades. IBGE. Coração de Maria; 2020.
- Lobo TCF. Para uma sociolinguística histórica do português no Brasil: edição filológica e análise linguística de cartas particulares do recôncavo da Bahia, século XIX [tese]. São Paulo: Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo; 2001.
- Lobo TCF. Arquivos, acervos e a reconstrução histórica do português brasileiro. In: Oliveira K, Souza HFC, Soledade J, organizadores. Do português arcaico ao português brasileiro: outras histórias. Salvador: EDUFBA; 2009.
- Lobo TCF. Rosa Virgínia Mattos e Silva e a história social linguística do Brasil. Estudos de linguística galega. 2015;7:69-82.
- Lose AD, et al. Edições de documentos históricos: a quem interessam? a quem se destinam? Revista da ABRALIN. 2017 jan.-abr.;16(2):71-86.
- Lucchesi D. As duas grandes vertentes da história sociolinguística do Brasil (1500-2000). Delta. 2001; 17(1):97-132.
- Lucchesi D. Língua e sociedade partidas: a polarização sociolinguística do Brasil. São Paulo: Contexto; 2015.
- Lucchesi D, Baxter A. A transmissão linguística irregular. In: Lucchesi D, Baxter A, Ribeiro I, organizadores. O português afro-brasileiro. Salvador: EDUFBA; 2009. p. 101-124.
- Maia C. Linguística histórica e filologia. In: Lobo T, et al., organizadores. Rosae: linguística histórica, história das línguas e outras histórias. Salvador: EDUFBA; 2012. p. 533-542.
- Mattos e Silva RV. Para a história do português culto e popular brasileiro: sugestões para uma pauta de pesquisa. In: Alkmim TM, organizadora. Para a história do português brasileiro. Vol. 1. São Paulo: Humanitas/FFLCH; 2002. p. 443-464.
- Mattos e Silva RV. Ensaios para uma sócio-história do português brasileiro. São Paulo: Parábola editorial; 2004.
- Mattos e Silva RV. Caminhos da linguística histórica: ouvir o inaudível. São Paulo: Parábola editorial; 2008. p. 7-26.
- Mussa A. O papel das línguas africanas na história do português do Brasil [dissertação]. Mimeo. Rio de Janeiro: UFRJ; 1991.
- Nascimento MFB. O lugar do corpus na investigação linguística. [citado 20 abr. 2017]. Disponível em: <http://www.clul.ul.pt/equipa/berlim-2000-nascimento.pdf>.
- Oliveira K. Negros e escrita na Bahia do século XIX: sócio-história, edição filológica de documentos e estudo lingüístico [tese]. Salvador: Universidade Federal da Bahia, Instituto de letras; 2006.
- Oliveira MF, Soares M, Silva AJ, organizadores [CD]. Correspondências amigas: o acervo de Valente, Bahia (1980-1993): edição fac-similar. Vol. 1. Feira de Santana: UEFS editora; 2011.

Petrucci A. Alfabetismo, escritura, sociedad. Barcelona: Gedisa; 1999.

Petrucci A. La ciencia de la escritura. Primera lección de paleografía. Buenos Aires: Fondo de cultura económica de Argentina; 2003

Queiroz RCR. Filologia e lexicologia: a edição e o estudo do vocabulário de autos de defloramento. Linguagem - estudos e pesquisas. 2012;16(2):15-28.

Santiago HS, organizador. [CD] Cartas em sisal: Riachão do Jacuípe, Conceição do Coité e Ichu (1906-2000): edição fac-similada. Vol. 3. In. Santiago HS, Carneiro ZON, Oliveira K, organizadores. Volume 3 de cartas brasileiras (1809-2000): coletânea de fontes para o estudo do português. Feira de Santana: UEFS editora; 2011.

Santiago HS. Um estudo do português popular brasileiro em cartas pessoais de mãos “cândidas” do sertão baiano [dissertação]. Feira de Santana: Universidade Estadual de Feira de Santana; 2012.

Sousa MCP, Kepler FN. E-dictor: uma ferramenta integrada para a anotação de edição e classe de palavras. VI encontro de linguística de corpus. São Paulo; 2007. [citado 26 abr. 2020]. Disponível em: https://www.ime.usp.br/~tycho/participants/psousa/edictor/presentation/edictor_2007.html.

Sousa MC, Kepler FN, Faria PPF. E-dictor: versão 1.0 beta 10, 2013 [programa de computador]. [citado 20 abr. 2020]. Disponível em: <http://edictor.net/download/>.

FLP 24(1)