Revista A Cor das Letras

Revista Digital dos Programas de
Pós-Graduação do Departamento de Letras e Artes da UEFS
Feira de Santana, v. 23, n. 2, p. 57-84, maio-agosto, 2022

http://periodicos.uefs.br/index.php/acordasletras/index
10.13102/cl.v23i2.7897
ISSN 2594-9675

De baús de famílias da região rural de Coração de Maria para o banco CE-DOHS: um acervo de cartas para a pesquisa em sociolinguística histórica[1]

From family trunks in the rural region of Coração de Maria to the CE-DOHS bank: a collection of letters for research in historical sociolinguistics

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

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

Resumo: Neste texto, apresentam-se algumas reflexões sobre o *corpus* Cartas Marienses, documentação epistolar de foro privado, escrita por trabalhadores rurais e donas de casa, ao longo do século XX, na região rural de Coração de Maria, interior baiano. A amostra disponibilizada em Brito (2020) atende a uma das agendas do Programa para História do Português Brasileiro (PHPB) – a constituição de *corpora* diacrônicos – e foi constituída no âmbito do projeto *Corpus* Eletrônico de Documentos Históricos do Sertão, parceiro do PHPB. As Cartas Marienses, escritas por indivíduos que tiveram pouco acesso às normas gramaticais disseminadas pelo processo de escolarização, por isso, pouco familiarizados com o código escrito, apresentam marcas de inabilidade em escrita alfabética em vários planos. O *Corpus* passou por uma refinada metodologia de contextualização sócio-histórica (PETRUCCI, 2003) e tratamento filológico, o qual uniu métodos da Antiga e da Nova Filologia, com disponibilização de edição semidiplomática e de edição modernizada, em linguagem XML, com uso do *eDictor*. A amostra em questão é de grande valia para os estudos histórico-diacrônicos do português brasileiro (PB) no âmbito da Linguística Histórica e, também, da História Social da Cultura Escrita, uma vez que visibiliza uma escrita cotidiana. Com pouco grau de monitoramento, a escrita revela traços de marcas orais, produto de *mãos pouco hábeis*, de pessoas comuns, de classes sociais desfavorecidas, que ficaram à margem da sociedade, mas que, nas cartas que escreveram, também contaram a sua história.

Palavras-chave: Banco de dados. Português Popular Brasileiro. Cultura Escrita.

Abstract: In this text, we present some reflections on the corpus Cartas Marienses, private epistolary documentation, written by rural workers and housewives, throughout the 20th century, in the rural region of Coração de Maria, in the interior of Bahia. The sample available in Brito (2020) meets one of the agendas of the Program for the History of Brazilian Portuguese (PHPB) – the constitution of diachronic corpora – and was constituted within the scope of the project Corpus Eletrônico de DocumentosHistóricos do Sertão, a

[1]Texto escrito como parte dos resultados da pesquisa de mestrado de Brito (2020), intitulada Cartas Marienses (Séc. XX): edição fac-similar e semidiplomática e estudo da concordância nominal, desenvolvida no âmbito do Programa de Pós-Graduação em Estudos Linguísticos (PPGEL), da Universidade Estadual de Feira de Santana, e orientada pela professora doutora Mariana Fagundes de Oliveira Lacerda.

*Estudante de doutorado na Universidade Estadual de Feira de Santana. E-mail: patysantosjesus@hotmail.com

*Doutora em Letras e Linguística e professora do Departamento de Letras e Artes, da Universidade Estadual de Feira de Santana. E-mail: marianafag@gmail.com

PHPB partner. The Cartas Marienses, written by individuals who had little access to the grammatical norms disseminated by the schooling process, therefore, little familiar with the written code, present marks of inability in alphabetical writing in several planes. The Corpus underwent a refined methodology of socio-historical contextualization (PETRUCCI, 2003) and philological treatment, which combined Old and New Philology methods, with the availability of semi-diplomatic and modernized editions, in XML language, using eDictor. The sample in question is of great value for the historical-diachronic studies of Brazilian Portuguese (BP) in the scope of Historical Linguistics and, also, of the Social History of Written Culture, since it makes visible an everyday writing. With little degree of monitoring, writing reveals traces of oral marks, the product of unskilled hands, common people, from disadvantaged social classes, who were on the margins of society, but who, in the letters they wrote, also told their story.

Keywords: Database. Brazilian Popular Portuguese. WrittenCulture.

1 PARA COMEÇAR

As fontes documentais remanescentes do passado têm servido de material empírico para diferentes investigações linguísticas, tanto no âmbito da Linguística Histórica, como também para a História Social da Cultura Escrita. É com o suporte dos *corpora* histórico-diacrônicos que a reconstrução do passado do português brasileiro tem avançado na direção de reconstruir estágios de língua não mais acessíveis, pois, como bem destacou outrora Mattos e Silva (2008, p. 21), são os dados empíricos fornecidos por documentos dessa natureza, e com o suporte da Linguística Histórica, que se podem chegar "à caracterização esquemática de um momento da história passada de uma língua" permitindo, por sua vez, "rastrear e em parte reconstruir seu uso vivo", levando-se em conta, a impossibilidade de acessar a língua falada de sincronias mais afastadas.

Por conseguinte, a constituição de *corpora* e/ou banco de dados sociolinguísticos histórico-diacrônicos configura-se em uma agenda primordial e necessária no âmbito do Projeto Nacional para a História do Português Brasileiro (PHPB). Os textos produzidos pelos segmentos marginalizados socialmente e que não tiveram acesso pleno à escolarização trazem, possivelmente, marcas do vernáculo e "permitem entreouvir a voz pelas frestas da variação da escrita" (MATTOS E SILVA, 2008, p. 21).

No campo da História Social da Cultura Escrita, Petrucci (1962) ressalta a necessidade de resgatar os materiais escritos marginalizados pela história, especialmente as escritas usuais, testemunhos que são mais difíceis de encontrar.

Barbosa (2008, p. 183) afirma, portanto, que é preciso descobrir fontes escritas mais transparentes, com índices que reflitam um pouco melhor certos traços de oralidade, seja pela inabilidade/desconhecimento por parte de seus redatores de fórmulas textuais, seja pela relação de simetria entre destinatário e remetente, seja ainda pelo caráter intimista de um dado tema. Oliveira (2009), por sua vez, observa a importância de constituir *corpora* que apresentem uma escrita mais próxima do vernáculo, sobretudo dos grupos sociais subalternos.

No âmbito da Linguística Histórica, Mattos e Silva (2004) afirma que, para compreender e interpretar efetivamente o português brasileiro, muitas histórias de falantes devem ser reconstruídas, e o caminho mais evidente é o de investigar fontes históricas de vária natureza para refazer a história das sucessivas distribuições demográfico-linguísticas

no espaço e na sociedade, em especial, o português popular, constituído, sobretudo na oralidade.

O *Corpus* Eletrônico de Documentos Históricos do Sertão (CE-DOHS), do Núcleo de Estudos de Língua Portuguesa (NELP) da Universidade Estadual de Feira de Santana (UEFS), Bahia, é um dos projetos parceiros do PHPB. É do banco CE-DOHS o acervo aqui considerado, as Cartas Marienses, documentação novecentista.

Cartas Marienses é um *corpus* ilustrativo das classes sociais desprestigiadas socialmente; está disponível, na plataforma CE-DOHS, nas versões fac-similar, semidiplomática e modernizada, para servir de material empírico, podendo contribuir com as investigações sobre o português brasileiro, especialmente sobre a penetração e difusão da língua escrita no interior baiano.

O material é um conjunto de 89 correspondências pessoais, entre cartões, cartas e bilhetes de foro privado, escritos entre o primeiro e último quartel do século XX, por trabalhadores rurais e donas de casa do interior da Bahia, pouco escolarizados e de classes sociais menos abastadas.

Essas correspondências foram escritas no cotidiano, sem muitas formalidades, trocadas por pessoas que mantinham uma relação de intimidade/proximidade, as quais evidenciam uma escrita transparente aos usos vernáculos, com traços da oralidade.

O presente texto, que tem como objetivo apresentar o acervo supracitado e seu valor para as pesquisas em Linguística Histórica e História Social da Cultura Escrita, está organizado da seguinte forma: em 1, as considerações iniciais; em 2, apresentam-se o banco de dados CE-DOHS e a metodologia adotada para o tratamento dos *corpora* disponibilizados na plataforma; em 3, apresentam-se o acervo e a contextualização sócio-histórica; em 4, reflete-se sobre o fazer filológico como curadoria do patrimônio histórico; em 5, discute-se sobre o fazer filológico frente às humanidades digitais; em 6, fala-se sobre a inabilidade em escrita alfabética e trazem-se alguns dados e reflexões sobre o *corpus*; em 7, as reflexões finais seguidas das referências consultadas.

2 CE-DOHS: MAIS DO QUE UM BANCO DE TEXTOS, UM BANCO DE DADOS SOCIOLINGUÍSTICOS

O projeto *Corpus* Eletrônico de Documentos Históricos do Sertão (CE-DOHS) resulta do desdobramento de outros dois projetos, o Projeto Contribuições para a constituição de um banco de textos e de um banco de dados para o estudo da história do português do Brasil, do séc. XVII ao XX, criado em 1997, e do Vozes do Sertão em Dados: história, povos e formação do português brasileiro criado em 2009, do qual faz parte o projeto DOHS: Documentos Históricos do Sertão. Ambos fazem parte do Núcleo de Estudos de Língua Portuguesa (NELP), da Universidade Estadual de Feira de Santana, maior núcleo de estudos de Língua portuguesa, em perspectiva histórica, da Bahia.

No que diz respeito à linguística de *corpus*, O CE-DOHS é pioneiro no Nordeste. Em 2010, passou a disponibilizar na plataforma digital um banco de dados sociolinguísticos que apresenta uma base documental, majoritariamente inédita, para estudo do português brasileiro, no âmbito da Linguística Histórica, da Linguística

Diacrônica, da História Social da Cultura Escrita e das demais disciplinas da Linguística, bem como áreas afins e outras áreas do conhecimento.

O referido projeto está organizado em duas fases, sendo que a primeira compõe-se de textos escritos entre 1750 a 2000, representativos do período histórico do PB, caracterizado pelo multilinguismo localizado. Esse banco disponibiliza textos do português culto, semiculto e popular, os quais permitem estudar a história do português brasileiro nesse contexto. Já a segunda fase, iniciada em 2019, tem por objetivo estudar a gestação do PB, por isso, volta-se para uma sincronia mais afastada, na qual o multilinguismo no Brasil era generalizado. Nesta fase, há uma escassez de fontes, considerando que são raros os textos produzidos por grupos nascidos no Brasil, sobretudo de índios e negros, etnias que não tiveram acesso à escola; há no projeto, porém, pequenos acervos bastantes significativos.

Para a composição do banco de textos no âmbito do projeto, exige-se um alinhamento teórico e metodológico como proposto pela sociolinguística histórica (ROMAINE, 1982; HERNÀNDEX-CAMPOY; CONDE-SILVESTRE, 2012). Na contextualização sócio-histórica, controlam-se informações quanto à autoria, à identificação da naturalidade e nacionalidade dos redatores, ao perfil social dos escreventes, se eram escolarizados, à profissão, às funções desempenadas, ao grau de habilidade com a escrita, qual etnia pertenciam, à idade, à cidade onde produziram os textos, entre outras informações.

Quanto ao tratamento metodológico dos textos, disponibiliza edições nas versões fac-similar, semidiplomática, diplomática e modernizada. As edições semidiplomáticas são feitas com fidedignidade e critérios bem definidos, elaborados com base nas normas de transcrição de documentos manuscritos do PHPB.

As edições modernizadas são disponibilizadas em linguagem XML *(eXtensible Markup Language)*, feitas por meio da ferramenta *eDictor*, um editor voltado para o tratamento filológico digital e à análise linguística automática, desenvolvido por Paixão de Souza, Kepler e Faria (2007, 2013). A referida ferramenta permite a geração automática de versões correspondentes às edições diplomáticas, semidiplomática e modernizadas, em versões HTML e XML com anotação morfossintática.

Figura 1 – *Corpus* Compartilhado Diacrônico.



Fonte: CE-DOHS (http://www.tycho.iel.unicamp.br/cedohs/corpora.html).

A base documental resulta de pesquisa em dezenas de arquivos públicos e privados, nacionais e internacionais – e também de gravações de fala, em diversas regiões da Bahia. Como se vê, um trabalho de pesquisa que envolve professores, estudantes de iniciação científica, mestrandos e doutorandos, além de pesquisadores colaboradores de outros projetos e instituições parceiras, em amplas frentes, a saber:

i. *prospecção e edição filológica*, em cooperação com o *Programa para a História do português (PROHPOR)*, coordenado por Rosa Virgínia Mattos e Silva, com contribuição direta ao seu *Banco Informatizado de Textos* (BIT), na inflexão do PB, desde 1994; a partir de 1997, contribuições ao *Projeto Nacional Para a História do Português Brasileiro* (PHPB), coordenado por Ataliba Castilho, no *âmbito do PHPB-Equipe Bahia (Salvador/Feira de Santana);*
ii. *edição digital*, em parceria com o *Corpus Histórico do Português Tycho Brahe (CTB)*, coordenado por Charlotte Marie Chambelland Galves, e *compartilhamento de ferramentas computacionais*, por meio de parcerias com outras instituições de Ensino Superior, em alguns casos, mediante acordo de cooperação formal entre a UEFS e a UNICAMP.
iii. *anotação sintática*, via coparticipação com o Projeto *CorPorA II: um corpus anotado para a história do português brasileiro*, que propõe a criação de uma plataforma comum (CNPq Edital Universal, em 2018), o chamado *Corpus* CTB-Brasil, mediante acordo de cooperação formal entre a UEFS e a Unicamp. CTB que, por sua vez, segue o padrão usado nos *Penn Parsed Corpora of Historical English*, também utilizado no *Corpus* de Francês do projeto *Modeliserle Changement: Lesvoies du français* e no *Corpus histórico do islandês*. Na mesma linha de outros *corpora* do português que vêm sendo anotados, o *Corpus* do português antigo, *WOCHWEL* , o *Corpus* dialetal *CORDIAL SIN* e o *Corpus* diacrônico de cartas *PS-Post Scriptum.* (CARNEIRO; LACERDA/CORPUS ELETRÔNICO DE DOCUMENTOS HISTÓRICOS DO SERTÃO, 2022.

Os *corpora* disponibilizados na Plataforma, cujo endereço é http://www5.uefs.br/cedohs/, foram produzidos por brasileiros de etnias, classes sociais e grau de escolaridade distintos, nascidos entre 1586 e 1986, e, também, por portugueses nascidos entre 1450 e 1850. Há no banco 50 acervos disponíveis, que somam cerca de um

milhão de palavras. Os *corpora* escritos somam 36 acervos, de gêneros textuais diversos, sendo que a documentação epistolar se sobressai, com mais de 1000 cartas pessoais.

No CE-DOHS,

> o pesquisador pode, acessando a plataforma, personalizar o *corpus*, de acordo com seu interesse; pode optar por separar os materiais, no que concerne ao autor, por: etnia (indígenas, brancos, negros do Brasil, mulatos, mamelucos e pardos); nível de escolarização e habilidade/inabilidade com a escrita; sexo; profissão; estratificação social; data e local de nascimento do autor (naturalidade e nacionalidade); no que concerne ao documento, pode separar materiais por: data e local de escrita, meio urbano e meio rural, para quê foi escrito e a quem foi destinado (SANTIAGO *et al*, p. 324-325, 2021).

Todos os materiais estão sendo, paulatinamente, disponibilizados na rede mundial de computadores, com acesso livre gratuito, para além de um acesso específico e parcial do *Corpus*, por meio de publicações. A coleção *Cartas brasileiras:* coletânea de fontes para o estudo do português (CARNEIRO, 2011), em três volumes, resulta do trabalho que vem sendo desenvolvido no âmbito do CE-DOHS. Para além dos três volumes citados, estão para sair em 2022, o volume 4, organizado por Brito e Lacerda, que traz as Cartas Marienses, aqui abordadas, e o volume 5, organizado por Carneiro, Tuy Batista e Almeida, que traz o acervo da Família Estrela Tuy.

3 DE PEDIDOS DE CASAMENTO A NARRATIVAS DE SÃO JOÃO E OUTRAS FESTAS: AS CARTAS MARIENSES (SÉCULO XX)

Diante das dificuldades de encontrar materiais significativos e representativos, Barbosa (2006, p.764-765), observa que se o que queremos encontrar são os materiais escritos mais transparentes aos usos gerais da fala, "talvez sejam mesmo as missivas trocadas em circulação privada (pessoais ou as da administração privada) o que mais se aproxime de uma escrita cotidiana, aquela que tem o limite possível do que a escrita transparece dados da oralidade, ocorrências de formas novas na sociedade".

No entanto, para a reunião de documentos dessa natureza, tanto para os estudos linguísticos sócio-históricos no campo da Linguística Histórica (MATTOS E SILVA, 2002, p.23), quanto para os estudos sob a égide da História Social da Cultura Escrita (PETRUCCI, 2003), faz-se necessária uma refinada metodologia para a averiguação de diferentes aspectos: *que*? *quando*? *onde*? *como*? *quem*? e *para que*?, para um controle metodológico confiável. Sendo assim, de acordo com a proposta de Petrucci (2003, p. 7-8), buscou-se controlar as seguintes questões:

> 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*? Conqué técnicas, com qué instrumentos, sobre qué materiales, según qué modelos fue escrito ese texto;
> v) *Quiénlo realizo*? 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.

Para Saéz e Castillo (2000), essa perspectiva de investigação permite:

> conocer las funciones atribuidas política y socialmente a los productos escritos: el prestigio social de los escribientes; el poder del escrito, ya sea el que pertenece a los que pose en la capacidad de escribir y la ejercen o el que ostentan sobre la escritura los aparatos políticos; los contextos de aprendizaje -desde la familia a la escuela- y los contenidos de la enseñanza de la escritura; la significación social de los maestros de escritura; los procesos de intermediación gráfica y la relación con las actividades de escritura por parte de los analfabetos; el status social de los alfabetizados; la necesidad social de aprender a escribir; o las prácticas concretas del escrito, tanto en sus usos activos (escritura) como pasivos (lectura) (SÁEZ Y CASTILLO, p. 28, 2000).

O acervo denominado *Cartas Marienses - BA (Séc. XX)* é um conjunto de 89 documentos epistolares, de foro privado, dos quais 69 cartas, 17 cartões e cinco bilhetes, produzidos entre 1935 e 1995. São 29 remetentes (15 mulheres e 14 homens), trabalhadores rurais e donas de casa, com pouco nível de escolaridade, nascidos entre o último quartel do século XIX e terceiro quartel do século XX, na região rural de Coração de Maria, interior baiano.

A documentação foi localizada em 2018. Boa parte do acervo – 39 manuscritos – estava arquivada em uma pasta tipo catálogo e depositada em um baú, sob os cuidados da Família Pacheco. Os demais manuscritos aqui apresentados foram "garimpados" em pequenos acervos familiares.

Figura 2 e 3 – Pasta catálogo e baú onde foram arquivados os manuscritos da Família Pacheco



Fonte: Brito (2020).

Os manuscritos encontram-se organizados por acervo, no âmbito do qual se acham ordenados por data cronológica de escrita. O Acervo Maria José Pacheco da Silva

(AMJP) é o mais representativo numericamente; além dos 39 manuscritos inicialmente localizados, fazem dele parte outros manuscritos produzidos por membros da família, totalizando 51 documentos.

Tabela 1 – Organização dos manuscritos por acervo

Acervos	Quantidade de manuscritos
Acervo Maria José Pacheco (AMJP)	51
Acervo José Figueiredo Miranda (AJFM)	3
Acervo José Mendes de Oliveira (AJMO)	5
Acervo Família Onofre (AFO)	3
Acervo Antônio Muricy de Oliveira (AAMO)	4
Acervo Valdemira Nogueira Martins Araújo (AVNMA)	17
Acervo Maria José Ribeiro Brandão (MJRB	5
Carta avulsa (Avulsa)	1
TOTAL	**89**

Fonte: Brito (2020).

Correspondências pessoais como as aqui apresentadas – escritas por indivíduos com pouca escolarização, de classes menos favorecidas socialmente, em regiões rurais mais afastadas – são escassas, uma vez que, ao longo do processo sócio-histórico de formação da sociedade brasileira, esses segmentos ficaram afastados da cultura letrada, escrevendo, portanto, esporadicamente. Localizar fontes de sincronias passadas, com tais características, apresenta singular relevância para os estudiosos em Linguística Histórica, haja vista os obstáculos com que o pesquisador se depara na busca por documentos dessa natureza, como a raridade, a dispersão arquivística e o mínimo quociente de durabilidade (PETRUCCI, 1999).

3.1 LOCALIZAÇÃO TEMPORAL

As correspondências que compõem o acervo *Cartas Marienses* foram escritas entre 1935 e 1995. Nas décadas de 30, 40, 50 e 60, há um número pequeno de cartas. A escassez de escolas na região pode ajudar a explicar a insuficiência de documentos representativos de classes sociais mais baixas, uma vez que esses segmentos não tinham recursos para estudar em outras cidades. Outra questão a considerar é o fato de que muitas pessoas, por necessidade, priorizavam o trabalho, sobretudo na lavoura. As crianças aprendiam apenas as primeiras letras, em ambientes extraescolares, geralmente em casa ou com alguma vizinha, não tendo acesso ao ensino formal. A maioria das cartas data das décadas de 70, 80 e 90, o que pode relacionar-se ao crescimento no processo de escolarização havido no Brasil, nesse período, como apontou Lucchesi (2015).

Das 89 correspondências, não há indicação de data de escrita em 52; foi, entretanto, possível inferi-la, graças às informações coletadas nas entrevistas, como também por meio do próprio conteúdo das cartas, o qual traz evidências que contextualizam aspectos da época.

As cartas sem indicação de data de escrita supostamente foram escritas a partir dos finais da década de 60 até meados da década de 90. A redatora Maria José Pacheco escreveu 28 dessas cartas, provavelmente entre os anos de 1984 e 1995.

3.2 LOCALIZAÇÃO ESPACIAL

Das 89 correspondências, há indicação de local de escrita em 31. Dessas, 21 foram escritas na zona rural do município de Coração de Maria; 1 foi escrita no povoado de Laranjeira, em Bento Simões, município de Irará; e 9 foram escritas em Salvador por marienses que estavam residindo na capital baiana.

Observa-se que a busca por melhores condições de vida impulsionou o êxodo rural em direção aos grandes centros urbanos, sobretudo a partir da década de 1930, como bem ressaltou Lucchesi (2015), quando, segundo o autor, implementa-se o processo de urbanização e industrialização da sociedade brasileira. Dentre aqueles que buscaram novas oportunidades em grandes centros urbanos, destaca-se José Mendes de Oliveira, que se mudou para Salvador, no ano de 1948, onde permaneceu por doze anos, voltando para sua terra natal, em 1960, para casar-se com Maria Cardoso Leal, a quem enviou algumas cartas. Outros que deixaram sua terra de origem foram: José Egídio Oliveira, Isabel Santos Lima, Doraci Cerqueira Martins e Zélia Nogueira Martins.

Com relação às correspondências que não tiveram seus locais de escrita especificados, foi possível inferi-los, especialmente pela relação com as demais correspondências, pelo conteúdo expresso, pelo local de residência dos redatores e pelas informações coletadas nas entrevistas. Desse conjunto de correspondências, fazem parte 33 cartas, escritas por Maria José Pacheco, moradora da Fazenda Água Verde, em Coração de Maria. A redatora, no total, escreveu 39 cartas. É dela a maior parte das cartas, enviadas, principalmente a seus filhos.

3.3 OS REMETENTES

As informações sobre os remetentes foram coletadas em entrevistas com familiares, conhecidos, bem como em documentos pessoais. Nas entrevistas realizadas, o objetivo foi fazer um levantamento do perfil sociocultural dos remetentes. Os dados levantados foram catalogados em fichas, seguindo o modelo utilizado por Carneiro (2005).

Essas "Marias" e esses "Josés" nasceram na região rural de Coração de Maria. Eram trabalhadores rurais e donas de casa que tiravam seus sustentos das diversas atividades desenvolvidas no campo, como a plantação de pequenas lavouras (o abacaxi, a mandioca, o fumo, o milho) e a criação de animais (galinhas, bois, carneiros, porcos), entre outras atividades.

Figura 4 – Maria José Pacheco, década de 1980. Fazenda Água verde, Coração de Maria, Bahia



Fonte: Brito (2020).

Destaca-se que a falta de recursos financeiros, a necessidade de trabalhar e a carência de escolas na região contribuíram para que esses indivíduos fossem privados de uma educação formal plena. Alguns deles, nas entrevistas feitas para construir o perfil sociocultural, confessaram angustiados que sonhavam em frequentar a escola e concluir a formação, mas as dificuldades financeiras enfrentadas pela família, os impediram de continuar os estudos, pois precisavam trabalhar para ajudar no sustento da casa. Ademais, não havia escolas de nível secundário nas proximidades, fato que dificultava ainda mais a vida daqueles indivíduos oriundos de classes sociais mais baixas, que dependiam de recursos para se deslocar para cidades circunvizinhas.

3.4 OS DESTINATÁRIOS

As correspondências do acervo *Cartas Marienses* foram enviadas a 27 destinatários; em sua maioria, pessoas que mantinham uma relação próxima com os remetentes, sendo que remetentes e destinatários eram todos da mesma localidade e compartilhavam o mesmo contexto sociocultural.

A maioria das correspondências do acervo foram enviadas por Maria José Pacheco, entre 1984[2] e 1995, principalmente para seus filhos, que receberam 34 das 39 cartas enviadas por ela: Edelzuita, Rute, Salvelina, Vilma e José Carlos. As 5 cartas restantes foram destinadas para a irmã Dade (Maria da Natividade Pacheco), para seus netos Raquel, Junior, Jessé, Maria Clara, e para sua amiga Conceição.

As missivas são escritas usuais, produzidas em contextos informais, no correr da vida diária, entre os afazeres domésticos e a lida dos trabalhos no campo. Foram trocadas em uma relação de proximidade (simetria) entre remetente e destinatário, e versam sobre o cotidiano dos interlocutores. Muitas abordam questões triviais do dia a dia dos

[2] Por inferência.

remetentes, algumas trazem pedidos de casamentos, desejam felicitações por uma data especial, dão as notícias da vida diária, ao mesmo tempo em que solicitam informações sobe os entes queridos, enfim.

Outro conjunto de 17 correspondências, composto por cartas e cartões, foi destinado à Valdemira Nogueira Martins, conhecida por Dona Vavá. São cartas enviadas por amigas, irmãs, primas, sobrinhas, parabenizando-a, desejando-lhe boas festas, dando-lhe notícias, expressando-lhe saudades, carinho e afeto.

Figura 5 – Fotografia de Valdemira Nogueira Martins na época das correspondências



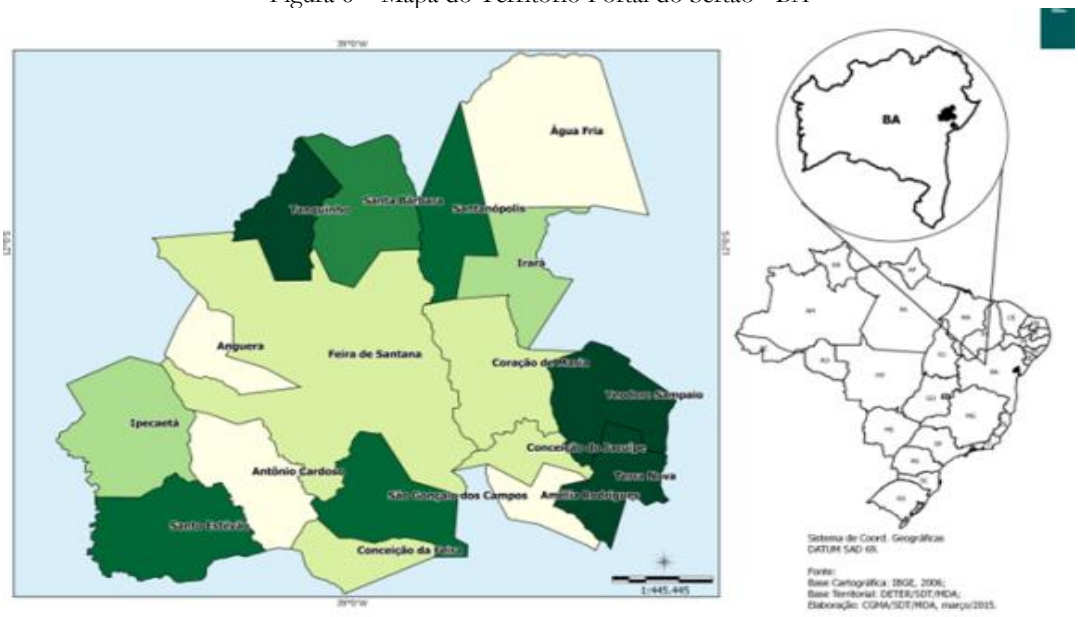
Fonte: Brito (2020).

Analice Vieira é a destinatária de 4 cartas, escritas e enviadas pelo seu marido, e que falam dos problemas cotidianos. As demais cartas foram destinadas a José Figueiredo, que recebeu 2 cartas, uma delas com um pedido de casamento; à Maria José, que recebeu 2 cartas; a José Mendes, que recebeu 2 cartas. Francisco, Maria, João Olympio, José Onofre, Maria José, Zilda, Tutu e Maria José Pacheco receberam 1 carta, cada um deles. Há, ainda, uma carta enviada a uma destinatária chamada "tia" e outra destinada 'a pais", ambas sem nomes específicos.

3.5 A REGIÃO: ASPECTOS SÓCIO-HISTÓRICOS

As correspondências foram produzidas por redatores oriundos do sertão baiano, especificamente por moradores da região rural do município de Coração de Maria. O município supracitado localiza-se na Zona Fisiográfica de Feira de Santana e está parcialmente incluído no Polígono das secas. De acordo com o Instituto Brasileiro de Geografia e Estatística (IBGE), possui uma área: 372, 315 km² 2 de extensão. Possui, segundo o Censo Demográfico realizado em 2010, 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. 58% se concentram na zona rural, e os outros 42%, na Zona urbana.

Figura 6 – Mapa do Território Portal do Sertão –BA



Fonte: http://sit.mda.gov.br/download/caderno/caderno_territoria.

Dados do IBGE indicam que o povoamento da localidade se iniciou na primeira metade do século XIX, quando da construção da Capela do Santíssimo Coração de Maria, no antigo povoado de Lages. No ano de 1853, o povoado foi elevado à condição de Freguesia e Distrito de Paz, passando a chamar-se Santíssimo Coração de Maria. Em 1891, o Arraial sede da Freguesia foi elevado à condição de Vila. A antiga Vila do Santíssimo Coração de Maria passou à categoria de cidade, em março de 1938, sendo extinta em 1943. Só em 1944, com o Decreto-Lei Estadual nº 12. 978, de 1º junho, o Município de Coração de Maria foi restabelecido, tendo como distritos o Arraial do Retiro e Itacava.

A sociedade Mariense é formada por descendentes de italianos e portugueses; embora não conste nos dados históricos do município, há indícios de que outros contingentes populacionais, a exemplo de indígenas, escravos e seus descendentes, habitaram a região.

Sobre a formação sociolinguística 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, de outro, os falares populares disseminados pela gente pobre que habitava a região.

Figura 7 – Plantação de Abacaxi, Zona Rural de Coração de Maria, 1992



Fonte: Brito (2020). Fotografia Cedida pelo Espaço Cultural Professora Rute Vieira , Fazenda Água Verde, Coração de Maria.

As principais atividades econômicas desenvolvidas pelos habitantes do município, desde o século XVIII, estavam voltadas para a agricultura e pecuária. Durante muito tempo, a plantação de pequenas lavouras e a criação de pequenos e grandes animais de corte movimentaram a economia local.

Em meados do século XX, o município é beneficiado pela política da Superintendência de Desenvolvimento do Nordeste (SUDENE). Os fazendeiros locais e pequenos produtores que já cultivavam, em pequena escala, a cultura do abacaxi; graças ao clima local favorável, ao solo fértil e à abundância de água, aproveitaram para buscar junto às entidades responsáveis, o financiamento para o cultivo, em maior escala, da lavoura.

A cultura do abacaxi, com o incentivo recebido, tornou-se a principal atividade econômica do município, elevando-o à condição de maior produtor do fruto no estado da Bahia, nas décadas de 70 e 80.

> Os moradores dessa jovem cidade, que driblavam as adversidades cotidianas, jamais poderiam imaginar que ela se tornaria a "terra do abacaxi" na década de 1970, responsável por 85% da produção baiana, suprindo praticamente todo o mercado consumidor do entorno e exportando o excedente para o sul do país, provocando mudanças significativas na vida dos seus moradores (SÃO PEDRO, 2018, p. 38).

No entanto, em meados da década de 80, houve um declínio na produção do fruto, sobretudo pela incidência da fusariose ou resinose, que afetou a lavoura, estagnando a economia local, que vivia, quase exclusivamente, dos rendimentos da plantação do abacaxi. Atualmente, muitos agricultores ainda produzem o fruto, em pequena escala, porém.

4 UMA VIAGEM ÀS DÉCADAS DE 30 A 90 DO SÉCULO XX: AS EDIÇÕES FAC-SIMILAR E SEMIDIPLOMÁTICA E A CURADORIA DO PATRIMÔNIO HISTÓRICO

Os *corpora* constituídos têm sido "a galinha dos ovos de ouro", no que diz respeito ao resgate e à reconstrução da história de uma língua, porquanto "só há história de uma língua se houver textos" (BARBOSA, 2008. p.181). Decerto, constituí-los não é uma tarefa simples ou fácil. Demanda do pesquisador amplos conhecimentos. Para além da história da língua, da materialidade linguística e tipologia textual, é fundamental possuir conhecimentos sobre tipo da escrita, morfologia das letras, suporte, instrumentos e técnicas utilizadas, entre outros. Por conseguinte, a Filologia, com seus refinamentos metodológicos, possibilita aos pesquisadores da área suporte para tal empreitada.

Em sentido estrito, a Filologia ou Filologia Textual se equipara à Crítica Textual, pois, segundo Malato (2017b. p. 60), é "a disciplina ou complexo de disciplinas destinadas à reconstrução e à interpretação, conduzidas com método científico dos textos literários e dos documentos linguísticos".

Por sua vez, em sentido amplo, a Filologia abarca

> [...] o complexo de estudos que, movendo-se em vários setores e utilizando-se de diversos instrumentos de investigação, baseando-se sempre em um exame crítico concreto dos textos, documentos e testemunhos, tem por objetivo uma exata e exauriente compreensão do próprio texto em sua precisa situação histórico-cultural e inclusive, em um escopo mais amplo, propõe-se o conhecimento integral e a reconstrução de um período histórico ou de uma ou mais civilizações, estudando a língua, a literatura, as diversas manifestações culturais (BALDUINO, 1979, p. 20).

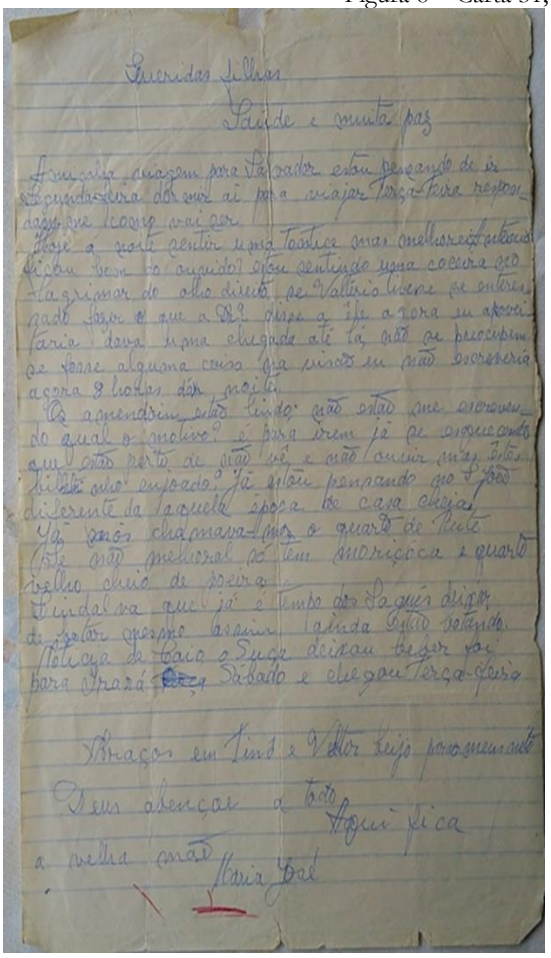
Aliada à Filologia, disciplinas auxiliares contribuem para as investigações, as quais descrevem e interpretam a dimensão material do texto, a exemplo da Codicologia, da Bibliografia material, da Ecdótica, da Manuscriptologia, da Paleografia em sentido amplo (a Paleografia de leitura, a Paleografia de análise e a História Social da Cultura Escrita), entre outras.

O *Corpus* Cartas Marienses, constituído no âmbito do CE-DOHS, resulta de pesquisa de mestrado desenvolvida por Brito (2020). Após o processo de localização, prospecção, seleção e organização documental, o referido acervo passou por uma refinada metodologia de tratamento filológico.

Inicialmente, a pesquisadora realizou a edição fac-similar, que configura a transferência para o meio digital, por meio de fotografia. Em seguida, fez a edição semidiplomática conservadora, que tem, como principais características, manter as propriedades linguísticas do documento e desenvolver abreviaturas. As conjecturas realizadas foram mínimas, apenas para situar o leitor. Para tanto, a edição seguiu critérios bem definidos e, com rigor filológico, elaborados com base nas normas do PHPB[3]. A seguir, fac-símile e edição semidiplomática da carta 31, do referido acervo.

[3] https://sites.google.com/site/corporaphpb/home/normas-de-edicao-do-phpb-2a-versao.

Figura 8 – Carta 31, Cartas Marienses



Carta 31

AMJP. Documento contendo um fólio. Papel almaço com pautas.

Queridas filhas|[4]

Saúde e muita paz|

A minha viagem para Salvador estou pensando de ir| Segunda-feira dormir ai para viajar Terça-Feira respon-| dam-me como vai ser |

Hoje á noite sentir uma tontice mas melhorei Antonio| ficou bem do ouvido? estou sentindo uma coceira no | lagrimar do olho direito se Valtério tivesse se enteres-|sado fazer o que a D*outo*ra disse a êle agora eu aprovei-|taria[5] dava uma chegada até lá, não se preocupem|se fosse alguma coisa na visão eu não escreveria | agora 8 horas dá~~r~~[6] noite

Os amendoin estão lindo não estão me escreven-|do qual o motivo? é para irem já se esquecendo | que estão perto de não vê e não ouvir mas êstes | bilhetinho[7] enjoado? Já estou pensando no S*ão* João |diferente da aquela época de casa cheia.|Já mós[8]chamava-mos o quarto de Rute.|

Se não melhoral só tem moriçoca e quarto | velho cheio de poeira |Lindalva que já é tempo dos Sagués deixar | de botar mesmo assim ainda estão botando|.

Noticia de Caio, o Suça deixou beber foi | para Irará ~~Terça~~[9] Sábado e chegou Terça-feira|

Abraços em Tinô e Valter[10] beijo para meus neto|

Deus abençoe a todo|

Aqui fica|

a velha mãe|

<u>Maria José</u>|

Fonte: Brito (2020). CE-DOHS.

[4] Escrita em 1992, na Fazenda Água Verde. (Por inferência).
[5] Rasurado.
[6] Rasurado.
[7] Rasurado.
[8] Rasurado.
[9] Rasurado.
[10] Rasurado.

5 DIÁLOGOS COM A NOVA FILOLOGIA: A EDIÇÃO EM LINGUAGEM XML

A partir da edição semidiplomática, o CE-DOHS desenvolve a edição modernizada[11], em linguagem XML (*eXtended Markup Language*), feita por meio do software *eDictor* (Paixão de Sousa e Kepler, 2007; Paixão de Sousa, Kepler e Faria, 2010), uma ferramenta desenvolvida especialmente para o tratamento de *corpora* digitais, que permite diferentes possibilidades de apresentação, edições semidiplomática, diplomática e modernizada.

A técnica moderniza a materialidade textual, padronizando grafias, acentuação e desenvolvendo abreviaturas, sem, contudo, perder informações genuínas ao documento, as quais podem ser recuperadas a contento do pesquisador, possibilitando o controle e mapeamento das intervenções realizadas nos textos. A versão digital mantém o rigor filológico, característico da edição semidiplomática, respeitando-se, todavia, as mudanças de parágrafo, de linha, os acidentes do suporte, as correções do autor, a orientação da escrita entre outros.

As versões modernizadas "atendem a diferentes objetivos: de um lado, visam a facilitar a busca textual por mecanismos automáticos; de outro lado, servem como base para ferramentas de anotação morfossintática no *Corpus* Tycho Brahe" (PAIXÃO DE SOUZA 2013, p. 125).

No que se refere à Linguística de *Corpus*, a ferramenta é um divisor de águas para o desenvolvimento de pesquisas, sendo a base da maioria dos *corpora* eletrônicos e edições filológicas digitais disponibilizadas atualmente. Como se vê:

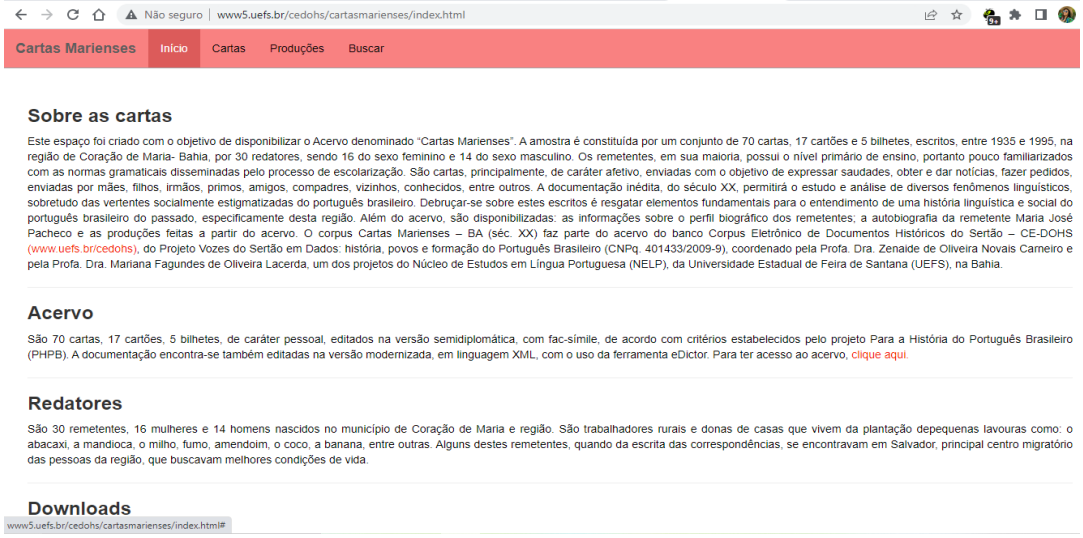
> Do feliz congraçamento entre as mais recentes tecnologias e a antiga Filologia, surgiu um novo universo de possibilidades para a preservação, disponibilização e análise de textos antigos, universo em que é possível oferecer ao leitor mais de uma edição do mesmo texto, permitindo que tenha a seu dispor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

Na era das humanidades digitais, o CE-DOHS, pioneiro do Nordeste em Linguística de *Corpus*, é um dos projetos a lançar mão dessa tecnologia.

As Cartas Marienses, cujas edições passaram pelo crivo da filologia, estão disponíveis para consulta e futuras pesquisas em diferentes versões de apresentação e edição no repositório do acervo: http://www5.uefs.br/cedohs/cartasmarienses/.

[11]A edição modernizada foi feita por Priscilla Tuy, pesquisadora do CE-DOHS.

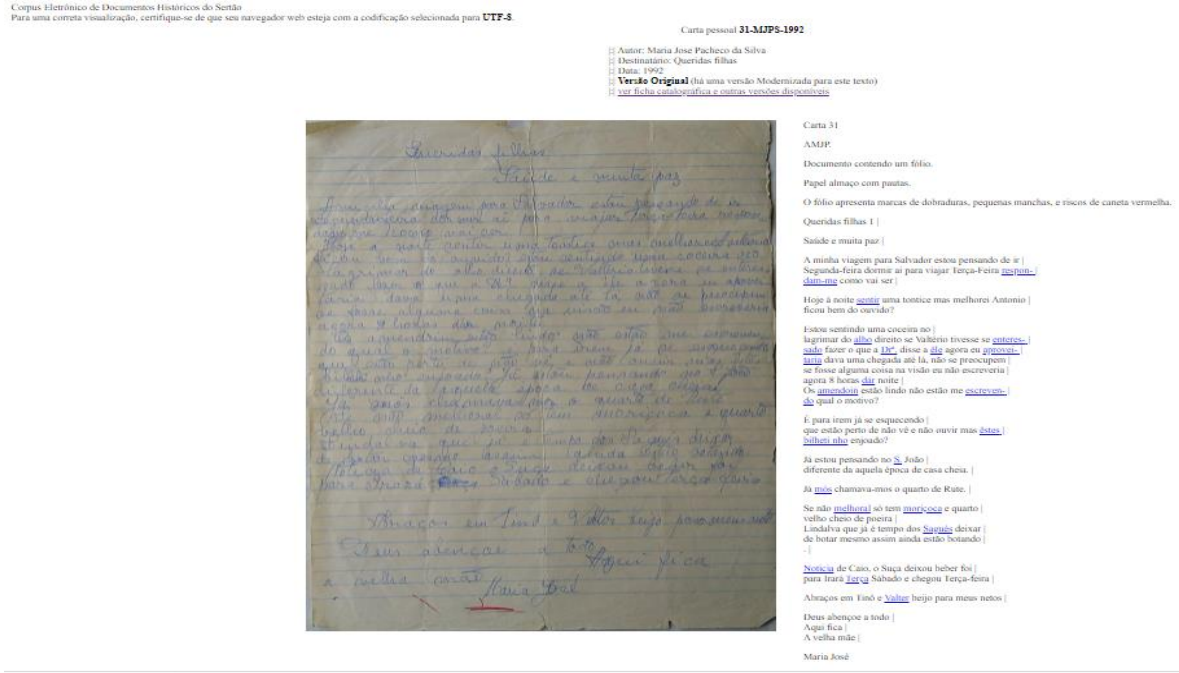
Figura 9 – Endereço eletrônico do acervo Cartas Marienses



Fonte: CE-DOHS.

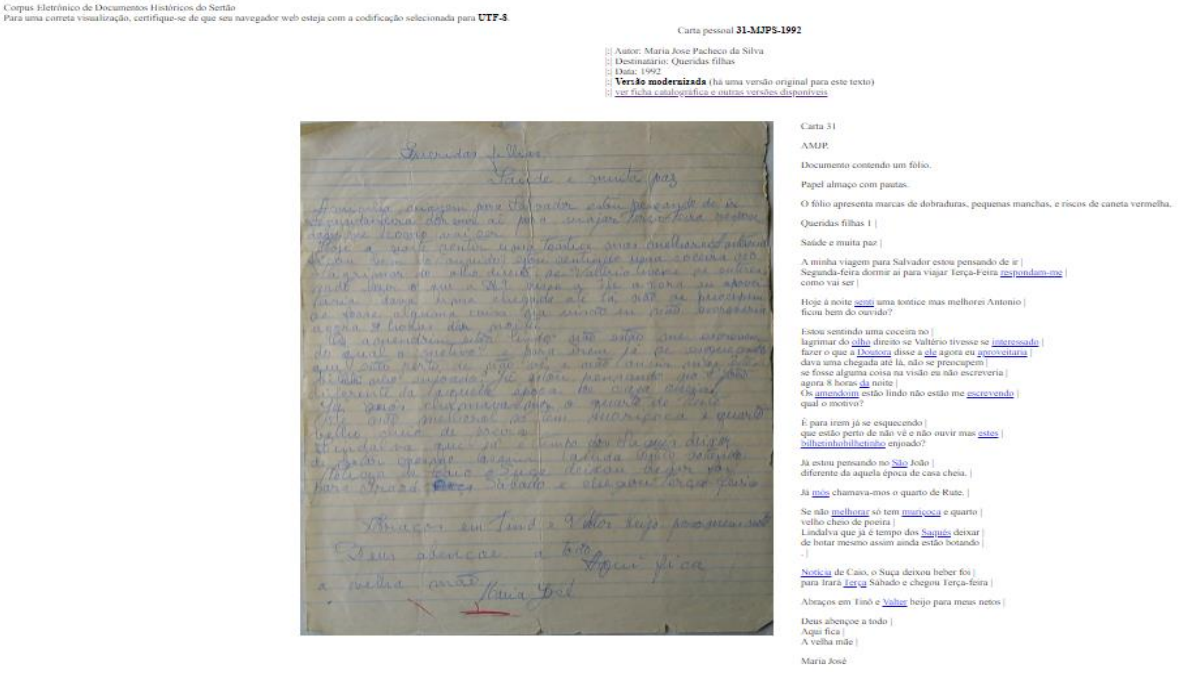
A seguir, uma ilustração da edição semidiplomática, modernizada final e da lista de dados modernizados, conforme apresentados na plataforma digital:

Figura 10 – Edição semidiplomática, com fac-símile, gerada no *eDictor*



Fonte: CE-DOHS.

Figura 11 – Edição modernizada final, com fac-símile, gerada no *eDictor*



Corpus Eletrônico de Documentos Históricos do Sertão
Para uma correta visualização, certifique-se de que seu navegador web esteja com a codificação selecionada para **UTF-8**.

Carta pessoal **31-MJPS-1992**

[-] Autor: Maria Jose Pacheco da Silva
[-] Destinatário: Queridas filhas
[-] Data: 1992
[-] **Versão modernizada** (há uma versão original para este texto)
[-] ver ficha catalográfica e outras versões disponíveis

Carta 31

AMJP.

Documento contendo um fólio.

Papel almaço com pautas.

O fólio apresenta marcas de dobraduras, pequenas manchas, e riscos de caneta vermelha.

Queridas filhas 1 |

Saúde e muita paz |

A minha viagem para Salvador estou pensando de ir |
Segunda-feira dormir ai para viajar Terça-Feira respondam-me |
como vai ser |

Hoje à noite sentí uma tontice mas melhorei Antonio |
ficou bem do ouvido?

Estou sentindo uma coceira no |
lagrimar do olho direito se Valtério tivesse se interessado |
fazer o que a Doutora disse a ele agora eu aproveitaria |
dava uma chegada até lá, não se precipitem |
se fosse alguma coisa na visão eu não escreveria |
agora 8 horas da noite |
Os amendoim estão lindo não estão me escrevendo |
qual o motivo?

É para irem já se esquecendo |
que estão perto de não vê e não ouvir mas estes |
bilhetinhobilhetinho enjoado?

Já estou pensando no São João |
diferente da aquela épica de casa cheia. |

Já mós chamava-mos o quarto de Rute. |

Se não melhorar só tem muriçoca e quarto |
velho cheio de poeira |
Lindalva que já é tempo dos Saqués deixar |
de botar mesmo assim ainda estão botando |
-|

Notícia de Caio, o Suça deixou beber foi |
para Irará Terça Sábado e chegou Terça-feira |

Abraços em Tinô e Valter beijo para meus netos |

Deus abençoe a todo |
Aqui fica |
A velha mãe |

Maria José

Fonte: CE-DOHS.

Figura 12 – Dados modernizados, gerados no *eDictor*



Carta pessoal **31-MJPS-1992**

[-] Autor: Maria Jose Pacheco da Silva
[-] Destinatário: Queridas filhas
[-] Data: 1992
[-] **Lista dos itens editados** (por ordem alfabética)
[-] ver ficha catalográfica e outras versões disponíveis

Item Original	Edições
alho	pad: olho,
amendoin	pad: amendoim,
aprovei- taria	seg: aproveitaria, jun: aprovei-taria,
bilheti nho	rasg: bilhetinho, jun: bilhetinho,
dár	pad: da, ras: r,
Drª.	exp: Doutora,
êle	pad: ele,
enteres- sado	pad: interessado, seg: enteressado, jun: enteres-sado,
escreven- do	seg: escrevendo, jun: escreven-do,
éstes	pad: estes,
melhoral	pad: melhorar,
moriçoca	pad: muriçoca,
mós	ras: mós,
Noticia	pad: Notícia,
respon- dam-me	seg: respondam-me, jun: respon-dam-me,
S.	exp: São,
Saqués	pad: Saqués,
sentir	pad: senti,
Terça	ras: Terça,
Valter	ras: Valter,

Fonte: CE-DOHS.

6 MARIAS E JOSÉS E AS MÃOS QUE ESCREVEM: DIMENSÕES DE INABILIDADE EM ESCRITA ALFABÉTICA

A inabilidade em escrita alfabética tem sido investigada em diferentes *corpora* de diferentes épocas e lugares. Marquilhas (2000) foi quem popularizou o termo *mão inábil* em língua portuguesa – utilizado anteriormente por Blanche-Benveniste (1993) – ao

investigar fatores externos à escrita em documentos da Inquisição portuguesa do século XVII, produzidos por indivíduos com poucas habilidades escriturais; no entanto, foi a partir de Petrucci (1973) que surgiram as bases para o desenvolvimento de tal conceito.

O termo *mão inábil*, do francês *scripteurs maladroits*, é uma metonímia da mão do redator pela escrita, que apresenta marcas características, comuns a indivíduos que pararam na fase inicial do processo de aquisição de escrita alfabética. Segundo Barbosa (2017), as *mãos inábeis* estão relacionadas às questões de *escriptualidade*, as quais revelam marcas de dificuldade de representação da escrita, sendo que a escrita fonética, com traços da oralidade, seria apenas mais um aspecto dessa dificuldade.

No âmbito do PHPB, algumas pesquisas que se debruçaram a investigar em *corpora* histórico-diacrônicos encontraram marcas de inabilidade em vários planos e dimensões. Cita-se o trabalho de Barbosa (1999), em cartas do comércio do século XVIII; Oliveira (2006), em atas escritas por africanos e afrodescendentes da Bahia do século XIX; e Santiago (2012), em cartas pessoais, escritas no século XX, por sertanejos baianos.

Para a identificação e reconhecimento dos textos das *mãos inábeis*, os pesquisadores buscam estabelecer alguns critérios que devem ser observados.

Marquilhas (2000) propõe observar características no plano físico-caligráfico, como traçado muito inseguro, incapacidade de alinhar perfeitamente as letras em um regramento ideal, uso de módulo grande, irregularidades de empaginação, entre outros, e; no plano da segmentação gráfica, no plano silábico e no plano infrassilábico.

Barbosa (2017, p. 22-28), com o objetivo de superar equívocos relacionados à identificação de um inábil, propõe, para o controle de marcas de inabilidade e identificação das *maõs inábeis*, a caracterização com base em nove dimensões, a saber: 1) da *escriptualidade* – os grafismos; 2) da escrita fonética – índices grafofonéticos; 3) da pontuação; 4) da repetição de vocábulos; 5) da dificuldade de riqueza na variação e precisão no léxico; 6) dos aspectos sintáticos; 7) das tendências discursivas; 8) da habilidade motora – níveis supragráficos e paleográficos; e 9. da segmentação gráfica – Hipersegmentação e Hipossegmentação.

Santiago (2019), com base nos critérios de Marquilhas (2000) e Barbosa (2017), apresenta uma possibilidade metodológica de reconhecimento de aspectos de inabilidade, a partir da caracterização dos perfis dos redatores, na qual busca verificar a coocorrência das dimensões acima referidas e a incidência maior ou menor de marcas em cada plano, para identificar em que ponto do contínuo a mão está situada: no ponto da *inabilidade máxima*; no ponto da *inabilidade parcial;* ou no ponto da *inabilidade mínima*.

No *corpus* Cartas Marienses, investigaram-se marcas de inabilidade em escrita alfabética, tomando-se, como base, os critérios estabelecidos por Barbosa (2017) e Santiago (2019); para tanto, selecionaram-se duas das nove dimensões propostas pelos referidos pesquisadores: da *escriptualidade*, grafismos, e da escrita fonética, os índices grafofonéticos.

6.1 ASPECTOS DE *ESCRIPTUALIDADE*

Alguns aspectos do código escrito, como a grafia de sílabas complexas, a representação da nasalidade, a grafação de dígrafos, entre outros, exigem do redator maiores habilidades de escrita; por conseguinte, indivíduos em estágios iniciais de aquisição de escrita alfabética não dominam plenamente tal tecnologia. Portanto, para a identificação das marcas de inabilidade na referida dimensão, Barbosa (2017, p. 24) ressalta que a investigação de tal fenômeno, em *corpora* histórico-diacrônicos, na prática, "[...] se traduz em observar inseguranças (ou em grau maior o quase desconhecimento do redator) em: (a) assumir sistematicamente grafismos (convencionalismo motivados por tradições culturais [...], (b) grafar sílaba complexa com /r/ ou /l/".

Indivíduos com poucas habilidades escriturais possuem dificuldades para grafar líquidas (/r/ ou /l/) em posição de coda ou em posição de ataque ramificado. Nestas construções, é comum que eles desfaçam a grafação CCV[12] ou CVC[13], com inclusão gráfica de vogal ou com a inversão de posição do /r/ ou /l/, explica Barbosa (2017). Nesses contextos de escrituração, a percepção do grau de inabilidade se apresenta de forma mais objetiva, especialmente quando os indivíduos não são inábeis em outros planos, como nos planos supragráficos e paleográficos reunidos.

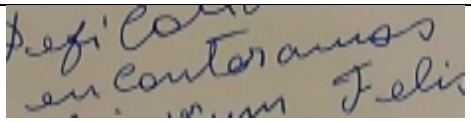
Nas Cartas Marienses, foram identificadas marcas de inabilidade relacionadas à grafação de sílabas complexas em <r> e <l>, tanto em posição de coda como de ataque ramificado, em situações de omissões e de deslocamentos. Com relação à nasalidade, destacam-se a representação exagerada em <n>~<mm>~<nn>, e também a ausência de representação em <Ø> ~ <m> ~ <n>.

Quanto aos aspectos relacionados aos grafismos, foram encontrados dados com representações de dígrafos com substituições de <rr> por <r>, de <ss> por <s>, de <ss> por <ç>, e de <ç> por <ss>, de <ch> por <x>, de <ch> por <sc>, e a omissão do <h>; na grafia de <lh> por <l>, de <nh> por <Ø >.

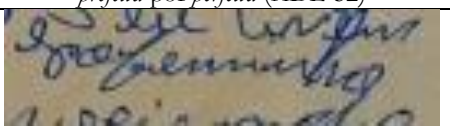
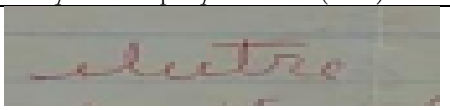
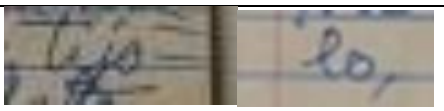
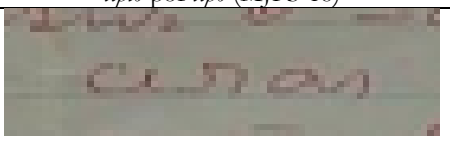
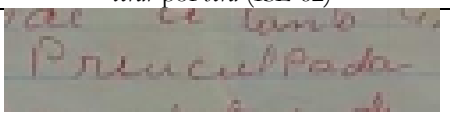
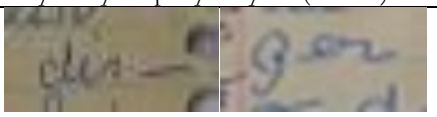
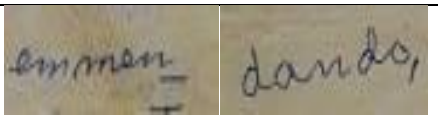
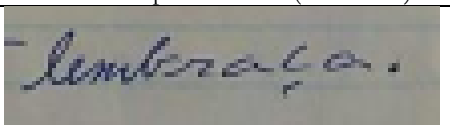
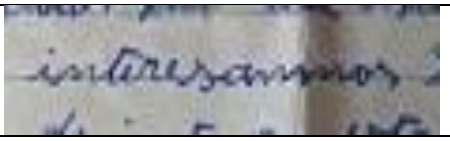
Os dados da tabela, a seguir, ilustram alguns aspectos encontrados no *corpus* Cartas Marienses.

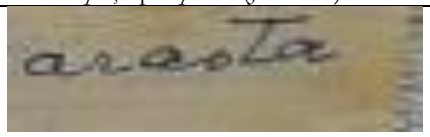
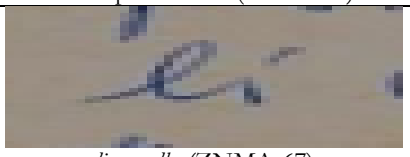
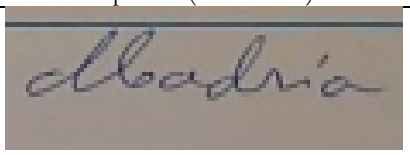
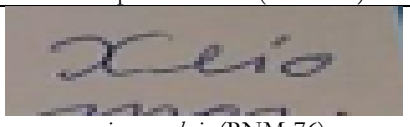
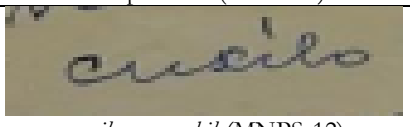
Tabela 2 – Dados de *escriptualidade* obtidos das Cartas Marienses

Grafia de sílabas complexas

Aspectos	Exemplos	Quantidade de ocorrências
Grafias com /r/ em ataque ramificado (deslocamento e omissões)	 encontaramos (MNM1-79)	3

[12](CCV) Consoante seguida de consoante e vogal.
[13](CVC) Consoante seguida de vogal e consoante.

Grafias com o /r/ em posição de coda (deslocamento e omissões)	 *prefeita* por *perfeita* (ADL-52)	10
Grafias com o /l/ (deslocamento e omissões)	 *pobeminha* por *probleminha* (E-89)	2
Grafias com o /s/ (omissões)	 *ilutre* por *ilustre* (ISL-62)	4
Hipercorreção		
Acréscimo de <l> em ataque ramificado	 *tiplo* por *tipo* (MJPS-18)	1
Acréscimo de <r> em posição de coda	 *cerar* por *será* (ISL-62)	16
Acréscimo de /l/ em posição de coda	 *preuculpada* por *preocupada* (ISL-62)	5
Acréscimo de <s> em posição de coda	 *diszer* por *dizer* (MJPS-21)	2
Representação da nasalidade		
Representação exagerada <n>~<mm>~<nn>	 *emmendando* por *emendando*(MNPS-12)	5
Ausência de representação <Ø> ~ <m> ~ <n>	 *lembraça* por *lembrança* (MNM2-80)	14
Representação de dígrafos		
<ss> ~ <s>		8

	interesammos por *interessamos* (EMO-58)	
<ç>~ <ss> ~ <s> ~ <ç>	 *preça* por *pressa* (JMO-56)	5
<rr> ~< r>	 *arasta* por *arrastar* (MNPS-12)	2
<lh> ~ <l>	 *li* por *lhe* (ZNMA-67)	2
<nh> ~ <Ø>	 *Madría* por *Madrinha* (RNM-76)	3
<ch> ~ <x>	 *xeio* por *cheio* (RNM-76)	1
<ch> ~ <sc>	 *cuscilo* por *cochilo*(MNPS-12)	1

Fonte: Elaborada pelas autoras.

6.2 ASPECTOS RELACIONADOS À ESCRITA FONÉTICA

Segundo Barbosa (2017), a inabilidade no nível fonológico de escrita diz respeito à estagnação do redator na fase de escrita fonética, da aquisição da escrita. Os indivíduos representam sons vocálicos e consonantais, na tentativa de imitar a pronúncia; tendem, pois, a se afastar das convenções gráficas.

A escrita mais fonética é um aspecto de dificuldade da representação da escrita alfabética, sendo a mesma em qualquer época no sentido de aquisição de escrita, porém, em razão das diferentes convenções ortográficas de cada época, é percebida de forma distinta nos *corpora* histórico-diacrônicos.

> No século XX, de convenção ortográfica homogeneizante, com uma grafia para cada palavra, qualquer dado grafo-fonético é denunciante de algum grau de escrita fonética. No século XIX, quando se valorizavam sobremaneira as grafias latinizantes e se retira a possibilidade de haver escrita fonética na ortografia, os dados grafo-fonéticos são marcas de inabilidade, mas não de forma tão contundente como no século XX (BARBOSA, p. 37. 2017).

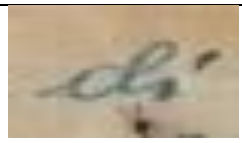
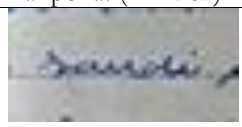
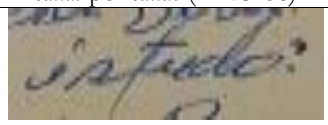
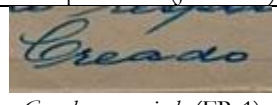
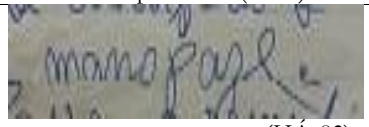
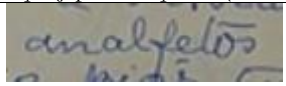
Em relação à análise de textos que apresentam esses índices, Santiago (2019, p.110) ressalta que "[...] é preciso considerar a variação gráfica normal ao período", pois a presença dessas marcas ou aparentes desvios podem não caracterizar uma mão inábil.

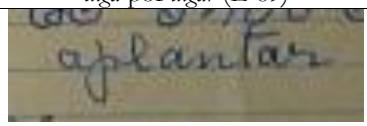
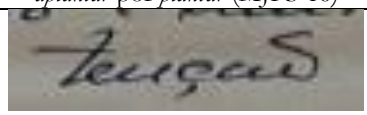
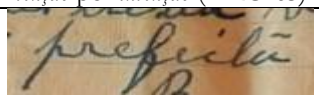
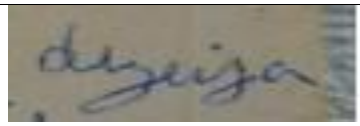
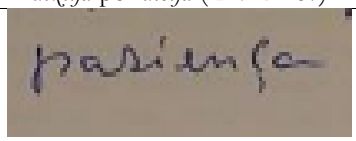
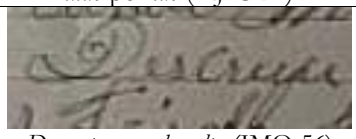
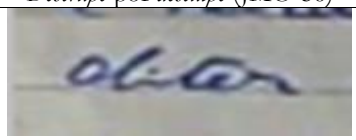
Nas Cartas Marienses, foram identificados dados referentes aos aspectos de elevação de vogais médias pretônicas [e] ~ [i] e [ẽ] ~ [ĩ] e [o] ~ [u] e [õ] ~ [ũ]; elevação de vogais médias postônicas [e] ~ [i] e [o] ~ [u]; elevação das vogais médias [e] ~ [i] e [ẽ] ~ [ĩ] e[o] ~ [u] e [õ] ~ [ũ] em monossílabos; abaixamento das vogais altas [i] ~ [e] e [ĩ] ~ [ẽ] e[u] ~ [o] e [ũ] ~ [õ] em posição pretônica; abaixamento de vogais em posição tônica; posteriorização de vogais; redução de ditongos orais e nasais; ditongação com a inserção da semivogal [y] e [w], nasalização, rotacismo, epêntese, síncope, apócope, prótese, aférese e metátese

Esses índices são mais gerais, comuns inclusive em mãos hábeis; em si, podem não constituir-se em marcas de inabilidade, mas são indícios que, ao lado de outros aspectos mais raros, podem refletir variantes diatópicas.

A tabela, a seguir, ilustra exemplos de índices grafofonéticos encontrados nas Cartas Marienses:

Tabela 3 – Dados de escrita fonética obtidos das Cartas Marienses

Aspectos	**Exemplos**	**Quantidade de ocorrências**
Elevação das vogais médias em monossílabos	 *di* por *de* (ADL-52)	37
Elevação de vogais médias postônica	 *saudi* por *saúde* (EMO-58)	15
Elevação de vogais médias pretônicas	 *istudo* por *estudo* (JCRB-85)	36
Abaixamento das vogais altas	 *Creado* por *criado* (FR-1)	47
Posteriorização de vogais	 *manopaze* por *menopausa* (HÁ-83)	3
Síncope	 *analfetos* por *analfabetos* (MJPS-47)	14

Apócope	 *luga* por *lugar* (E-89)	25
Prótese	 *aplantar* por *plantar* (MJPS-16)	14
Aférese	 *tenção* por *intenção* (AMO-63)	7
Metátese	 *prefeita* por *perfeita* (ADL-52)	2
Ditongação	 *dezeija* por *deseja* (ZNMA-67)	20
Redução de ditongos	 *pasiença* por *paciência* (MNM-71)	10
Nasalização	 *indo* por *ido* (MJPS-21)	10
Rotacismo	 *Descrupe* por *desculpe* (JMO-56)	2
Epêntese	 *Obiter* por *obter* (EMO-58)	1

Fonte: Elaborada pelas autoras.

Os dados apresentados, tanto no tópico 4.1, referente à dimensão da *escriptualidade*, quanto no tópico 4.2, referente à dimensão da escrita fonética, fornecem indícios de que esses redatores não dominam plenamente a tecnologia do código escrito, sendo, portanto, pouco familiarizados com o processo de escrita alfabética, adquirida, especialmente, por meio do ensino formal. Tais propriedades identificadas nas referidas dimensões, as quais

foram analisadas em coocorrência, trazem evidências de que os redatores são parcialmente inábeis em maior ou menor grau ou pouco hábeis na escrita.

7 CONCLUSÕES

Os *corpora* histórico-diacrônicos têm sido fundamentais para as pesquisas linguístico-históricas que, em diferentes frentes de trabalho e quadros teóricos específicos, buscam compreender e, em certa medida, reconstruir a sócio-história do português brasileiro, em especial no tocante à penetração e à difusão do português escrito, nos diversos cantos do Brasil. Na era das humanidades digitais, o CE-DOHS, pioneiro do Nordeste, oferece aos pesquisadores um amplo repositório documental, seriado controlado sócio historicamente.

O referido banco de dados conta com uma criteriosa contextualização sócio-histórica, que busca identificar uma série de informações referentes aos documentos e aos indivíduos que os produziram, como: o *que* escreveu? *quando* escreveu? *onde* escreveu? *como* escreveu? *quem* escreveu? *para quem* o texto foi escrito? e *para que* o texto foi escrito?, as quais asseguram confiabilidade à documentação.

No que tange ao tratamento filológico, o projeto CE-DOHS, ao qual este trabalho se vincula, conta com uma refinada metodologia, que agrega métodos da Antiga e da Nova Filologia, as quais permitem a preparação, apresentação e disponibilização de textos editados em diferentes versões.

O acervo aqui considerado, de Coração de Maria, dá a conhecer ao grande público vislumbres da vida corriqueira do sertão baiano ao mesmo tempo em que oferece, como se viu, um material rico às pesquisas em sociolinguística histórica.

REFERÊNCIAS

BALDUINO, A. *Manualedi filologia italiana*. Firenze: Sansoni, 1979.

BARBOSA, Afrânio Gonçalves. *Para uma história do português colonial:* aspectos linguísticos em cartas do comércio. Tese (Doutorado em Letras Vernáculas) – Universidade Federal do Rio de Janeiro, Rio de Janeiro, 1999, p. 134-204.

BARBOSA, Afrânio Gonçalves. Fontes escritas e história da língua portuguesa no Brasil: as cartas de comércio no século XVIII. In: LIMA, Ivana Stolze; CARMO, Laura do (Org.). *História social da língua nacional*. Rio de Janeiro: Edições Casa de Rui Barbosa, p. 181- 211, 2008. Disponívelem<http://www.casaruibarbosa.gov.br/dados/DOC/artigos/aj/FCRB_Historia_social_da_lingua_nacional.pdf >. Acesso em: 20/10/ 2020

BARBOSA, Afrânio Gonçalves. O controle de marcas de inabilidade na escrita alfabética e a identificação das mãos inábeis em corpora histórico-diacrônicos. *Revista da ABRALIN*, v.16, n.2, p. 19-43, Jan./Fev./Mar./Abril de 2017. Disponível em: <http://revistas.ufpr.br/abralin/article/view/51997/32039>. Acesso: 22/05/2018.

BLANCHE-BENVENISTE, Claire. Lesunités: langue écrite, langue orale. In: PONTECORVO, Clotilde; BLANCHE-BENVENISTE, Claire (Ed.). *Proceedings of the workshop onOrality versus Literacy*: concepts, methodsand data. Siena, Italy, 24-26 September 1992. Estrasburgo: Science European Foundation, 1993. p. 133-194.

BRITO, Patrícia Santos de Jesus. *Cartas Marienses (Séc. XX):* Edição semidiplomática e fac-similar e estudo da concordância nominal. Dissertação (Mestrado em Linguística) – Universidade Estadual de Feira de Santana, 2020.

BRITO, Patrícia Santos de Jesus; LACERDA, Mariana Fagundes de Oliveira (Org). *Cartas Brasileiras (1809-2000)*: coletânea de fontes para o estudo do português, volume 4. Feira de Santana. UEFS Editora, 2022.

CARNEIRO, Zenaide de Oliveira Novais. *Cartas brasileiras (1808-1904):* um estudo lingüístico-filológico Tese (Doutorado em Linguística) – Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, Campinas, São Paulo, 2005.

CARNEIRO, Zenaide de Oliveira Novais; LACERDA, Mariana Fagundes de Oliveira (Org). CE-DOHS - Corpus Eletrônico de Documentos Históricos do Sertão (2012-2025). Disponível em: www.uefs.br/cedohs. Acesso em: 26/04/2020.

CARNEIRO, Zenaide de Oliveira Novais; LACERDA, Mariana Fagundes de Oliveira (Org). CE-DOHS - Corpus Eletrônico de Documentos Históricos do Sertão (2012-2025). URL: http://www.uefs.br/cedohs. Accessed in: Acesso em: 20/02/2022.

GONÇALVES, Maria Filomena; BANZA, Ana Paula. Fontes de metalinguísticas para a história do português clássico. In: GONÇALVES, Maria Filomena; BANZA, Ana Paula. *Património Textual e Humanidades Digitais*: da antiga à nova filologia. Évora: CIDEHUS, 2013. p. 3-9

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

Instituto Brasileiro de Geografia e Estatística (IBGE). Coração de Maria-Bahia. Disponível em: https://cidades.ibge.gov.br/brasil/ba/coração demaria/historico. Acesso em: 27/03/2019.

LUCCHESI, Dante. *Língua e sociedade partidas*: a polarização sociolinguística do Brasil. São Paulo: Contexto, 2015.

MALATO, E. *Lessicofilollogico: unapproccioalla filologia*. Roma: Salerno, 2017b.

MARQUILHAS, Rita. *A faculdade das Letras:* leitura e escrita em Portugal no séc. XVII. Lisboa: Impressa Nacional-Casa da Moeda, 2000.

MATTOS E SILVA, Rosa Virgínia. *Ensaios para uma sócio-história do português brasileiro*. São Paulo: Parábola, 2004.

MATTOS E SILVA, RV. *Caminhos da linguística histórica:* ouvir o inaudível. São Paulo: Parábola editorial, 2008. p.7-26.

MATTOS E SILVA RV. Para a história do português culto e popular brasileiro: sugestões para uma pauta de pesquisa. In: AlkmimTM (Org.). *Para a história do português brasileiro.*v. 1. São Paulo: Humanitas/FFLCH, 2002a.p. 443-464.

OLIVEIRA K. *Negros e escrita na Bahia do século XIX:* sócio-história, edição filológica de documentos e estudo lingüístico. Tese (doutorado em letras e linguística) –Universidade Federal da Bahia, Instituto de letras, Salvador, 2006.

OLIVEIRA, Klebson. Ajuntamento de fontes para a história do português popular brasileiro: amores, desamores e outras espécies de dores. In: OLIVEIRA, Klebson; SOUZA, Hirão F. Cunha; GOMES, Luís. (Org.). *Novos tons de Rosa... para Rosa Virgínia Mattos e Silva.* Salvador: EDUFBA, 2009a.p. 175-195.

PAIXÃO DE SOUSA MC; keplerFN. *E-dictor: uma ferramenta integrada para a anotação de edição e classe de palavras*. VI encontro de linguística de *corpus*, São Paulo, 2007. Disponível em:<https://www.ime.usp.br/~tycho/participants/psousa/edictor/presentation/edictor_2007.html>.Acesso em: 26. 04. 2020.

PAIXÃO DE SOUSA MC ; keplerFN; Faria, PPF. *E-dictor*. versão 1.0 beta 10, 2013. *Programa de computador*. Disponível em: <http://edictor.net/download/>. acesso em: 20.04.2020.

PAIXÃO DE SOUZA. Maria Clara. A filologia Digital em Língua Portuguesa: alguns caminhos.In: GONÇALVES, Maria Filomena; BANZA, Ana Paula. *Património Textual e Humanidades Digitais*: da antiga à nova filologia. Évora: CIDEHUS, 2013. p. 113-138.

PETRUCCI, Armando. *Alfabetismo, escritura, sociedad*. Barcelona: Gedisa, 1999.

PETRUCCI, Armando , *Inlroduzione alie pratichediscrittura*, "AnnalidellaScuoia NórmaleSuperioredi Pisa", serie III, vol. XXIII-2 (1993) pp. 549-562.

PETRUCCI, Armando. *La ciencia de la escritura*: primeralección de paleografía. Buenos Aires: Fondo de Cultura Económica de Argentina, 2003.

PETRUCCI, Armando. Scrittura, alfabetismo ededucazionegraficanella Roma del primo

cinquecento: da um librettodicontidiMaddalenaPizzicarola in Trastevere. *Scrittura e Civiltá*, Roma, n. 3. p. 163-207, 1973.

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

SÁEZ, Carlos. E. Antonio. CASTILLO. 2000. "Paleografía e Historia de la Cultura Escrita: del signo a lo escrito", en*Introducción a laPaleografía y la Diplomática General*, editado por A. RiescoTerrero, 21-31. Madrid: Editorial Síntesis.

SANTIAGO, Huda Silva. *Um estudo do português popular brasileiro em cartas pessoais de mãos "cândidas" do sertão baiano*. Dissertação (Mestrado em Linguística) – Universidade Estadual de Feira de Santana, Feira de Santana, 2012.

SANTIAGO, Huda da Silva. *A escrita por "mãos inábeis*": uma proposta de caracterização. Huda da Silva Santiago. -- Salvador, 2019. Tese (Doutorado - Língua e Cultura) -- Universidade Federal da Bahia, Programa de Pós-Graduação em Língua e Cultura, 2019.

SANTIAGO, Huda da Silva; LACERDA, Mariana Fagundes de Oliveira; BRITO, Rosana Carvalho; CARNEIRO, Zenaide de Oliveira Novais. CE-DOHS: um banco de dados sociolinguísticos para a história do português brasileiro. *Labor Histórico*. [online]nº 7 (Especial): 311-329, 2021. Rio de Janeiro. Disponível em LaborHistórico (ufrj.br). Acesso em :09/03/2022.

SÃO PEDRO, Graziele Maia. *Frutos Dourados de Hastes Espinhosas*: Experiências na Produção de Abacaxis em Coração de Maria-Bahia (1960-1985). Dissertação (Mestrado em História) Universidade do Estado da Bahia, Campus II, 2018.