



UNIVERSIDADE ESTADUAL DE FEIRA DE SANTANA

Autorizada pelo Decreto Federal nº 77.496 de 27/04/76
Recredenciamento pelo Decreto nº 17.228 de 25/11/2016



PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
COORDENAÇÃO DE INICIAÇÃO CIENTÍFICA

XXIV SEMINÁRIO DE INICIAÇÃO CIENTÍFICA DA UEFS SEMANA NACIONAL DE CIÊNCIA E TECNOLOGIA - 2020

REVISÃO E ORGANIZAÇÃO DA EDIÇÃO DIGITAL DA DOCUMENTAÇÃO ESCRITA POR MÃOS INÁBEIS

1. Bolsista Larissa Nascimento Pedreira de Souza, PROBIC/UEFS, Graduanda em Licenciatura em Letras com Francês, Universidade Estadual de Feira de Santana, e-mail: ramonlarissa0226@outlook.com
2. Orientadora Huda da Silva Santiago, Departamento Letras e Artes, Universidade Estadual de Feira de Santana, e-mail: huda_santiago@uefs.br

PALAVRAS-CHAVE: *corpora* diacrônicos; *mãos inábeis*; edição digital.

INTRODUÇÃO

Os documentos escritos representam uma possibilidade de aproximação à oralidade, diante da impossibilidade de acesso direto à língua de períodos anteriores. A Linguística Histórica e a Filologia são responsáveis por estudar o desenvolvimento histórico da língua, levando em consideração as mudanças sociais que também acompanham esse processo, através desses documentos. Então, por causa da importância dessas fontes escritas, existe a necessidade de disponibilização e facilitação do acesso aos textos para as variadas finalidades, seja para pesquisa linguística ou para outras áreas. Isso se torna possível através dos diversos formatos de edição, já que, segundo Gonçalves e Banza (2013, p. 4) “surgiu um novo universo de possibilidades para a preservação, disponibilização e análise de textos antigos, universo em que é possível oferecer ao leitor mais de uma edição do mesmo texto, permitindo que tenha ao seu dispor o texto editado, em diferentes versões, e o seu original”; sendo agora disponibilizados em versão digital e eletrônica.

Nesse sentido, o Projeto Corpus Eletrônico de Documentos Históricos do Sertão (CE-DOHS) (Fapesb 5566/2010 - Consepe 202/2010), do Departamento de Letras e Artes (DLA) da Universidade Estadual de Feira de Santana (UEFS), colabora com a tarefa de constituição de *corpora* históricos, não literários, disponibilizando acervos que possuem textos mais próximos da escrita cotidiana, que reflitam certos traços de oralidade, seja pela inabilidade com a escrita, por parte dos redatores, seja pela relação simétrica entre redator e destinatário, como é o caso do acervo *Cartas em Sisal*, utilizado neste estudo.

O objetivo geral deste trabalho foi revisar a edição digital-eletrônica, modernizada, assim como atualizar os metadados do acervo *Cartas em Sisal*, constituído por 131 cartas escritas ao longo do século XX, no semiárido baiano, por remetentes que são pouco escolarizados, caracterizados como *mãos inábeis* em Santiago (2019), pois estagnaram na fase inicial da aquisição da escrita.

Esse trabalho de revisão das cartas e atualização das informações das fichas catalográficas (metadados do acervo) é importante para atender às necessidades dos estudos desenvolvidos pela equipe do projeto e para facilitar o acesso, aos demais

pesquisadores, aos textos que estão disponíveis em um site específico (www5.uefs.br/cedohs/maosinabeis/), vinculado ao site do CE-DOHS (<http://www5.uefs.br/cedohs/>), onde se disponibiliza, além da documentação escrita por *mãos inábeis*, editada nas versões fac-similar, semidiplomática e modernizada, as fichas catalográficas e todos os trabalhos produzidos a partir desse acervo. Como as cartas já haviam passado por um processo de revisão em trabalho anterior, os maiores necessidades de atualização foram referentes às fichas catalográficas (metadados).

MATERIAL E MÉTODOS

Para a revisão da edição e das fichas catalográficas do acervo, em linguagem XML, tendo como ponto de partida a edição semidiplomática e as fichas dos redatores publicadas em Santiago (2019), foi usada a ferramenta eDictor. O eDictor, criado por Paixão de Sousa, Kepler e Faria (2009), é uma ferramenta de anotação morfológica, sintática e edição eletrônica, seja edição semidiplomática, modernizada e/ou fac-símile, de textos em XML para finalidades de análise linguística. Essa otimização, através dessa ferramenta, contempla o trabalho e as análises desenvolvidas no Projeto CE-DOHS, que é um *corpus* que possui textos escritos por pessoas nascidas no Brasil, editados em XML. Neste projeto, o acervo Cartas em Sisal possui 131 cartas pessoais escritas por 31 mulheres e 22 homens, totalizando 53 remetentes que são pouco escolarizados, caracterizados como “mãos inábeis” (Santiago, 2019).

A fim de se obter dados significantes, para o estudo linguístico sócio-histórico é necessário considerar as informações sobre *quem* escreveu o texto e por qual causa, e não somente sobre o que está escrito, através da transcrição do texto. “Em relação aos estudos sócio-históricos linguísticos, com o objetivo de tornar melhores os “maus dados”, a que se refere Labov, geralmente tem sido possível indicar, com precisão, o *quando*, o *onde*, embora não se possa recuperar, na maioria das vezes, o perfil biográfico, tanto do emissor (*quem*), quanto do receptor (*para quem*)”, como comenta Mattos e Silva (2002). Para a autora, é muito relevante identificar quando, onde, quem escreveu e para quem foram escritos os textos. Devido à importância e necessidade dessa identificação, as fichas catalográficas dos textos do acervo *Cartas em Sisal* passaram por revisão e organização.

RESULTADOS E DISCUSSÃO

As fichas catalográficas do acervo contêm dados pessoais sobre os redatores e sobre as suas respectivas cartas, que são os dados do documento, e também as informações de processamento, que abrangem as etapas e os pesquisadores envolvidos no trabalho com o acervo. Antes da revisão, os itens desses metadados não seguiam uma ordem padrão, não possuíam letras maiúsculas no início, e faltavam alguns itens que deveriam ser contemplados. A Figura 1, a seguir, mostra o quadro que foi construído com sugestões de vários dos pesquisadores que integram o CE-DOHS, a partir das necessidades das pesquisas que têm sido desenvolvidas com base nos acervos, apresentando as mudanças necessárias para os metadados no “arquivo de configurações”. Também acompanham esse quadro as dicas e informações para quem for “gerar” os textos no sistema, ajustar as alterações propostas, como:

- retirar a pontuação final;
- inserir hífen para indicar que o campo não tem valor;
- colocar todas as datas no mesmo padrão;
- colocar todos os nomes conforme a ficha;
- não usar abreviaturas para os nomes de estados, cidades e países.

DADOS DO DOCUMENTO		
Conteúdo	Valor	Comentário
Acervo	Sigla	Conferir a sigla usada no acervo
Autor	Nome conforme ficha	Não usar o nome com abreviatura ou apelido. Usar o nome completo conforme a ficha
Sexo	Masculino/Feminino	Antes "gênero do autor" (excluir a informação antiga)
Data de nascimento	DD/MM/AAAA	
Nasido(N)/radicado(R)	Local de nascimento / Local que foi radicado	Conforme ficha
País de Nascimento	Conforme ficha	Manter sempre o nome completo do país, sem "apelidos", abreviaturas.
Cidade de Nascimento	Conforme ficha	Manter sempre o nome completo do estado, sem "apelidos", abreviaturas.
Estado de Nascimento	Conforme ficha	Manter sempre o nome completo da cidade, sem "apelidos", abreviaturas.
Escolaridade	-	Ainda não existe um padrão
Habilidade de Escrita	Hábil / Pouco Hábil / Inábil	Apenas no acervo de Huda (Inâbeis/Cândidas/Sisal)
Profissão	Conforme ficha	Manter sempre o maior grupo da profissão: Advogado Trabalhista. Profissão: Advogado
Local de Escrita	Conforme ficha	Preencher apenas se o local for diferente da cidade/estado. Ex.: Fazenda Riachão
País de Escrita	Conforme ficha	Manter sempre o nome completo do país, sem "apelidos", abreviaturas.
Estado de Escrita	Conforme ficha	Manter sempre o nome completo do estado, sem "apelidos", abreviaturas.
Cidade de Escrita	Conforme ficha	Manter sempre o nome completo da cidade, sem "apelidos", abreviaturas.
Data do Documento	DD/MM/AAAA	Data da escrita do documento
Destinatário	Nome conforme ficha	Não usar o nome com abreviatura ou apelido. Usar o nome completo conforme a ficha
Gênero do Documento	Carta / Oral / Impresso	
Relação Social	Amigo / Primo / Irmão	Conforme ficha
Relação Estabelecida	-	Ainda não existe um padrão
Conteúdo	Conforme ficha	
Fonte	Conforme acervo	Usar sempre o nome completo, sem abreviatura
Referência	Conforme acervo	Usar sempre o nome completo, sem abreviatura
PROCESSAMENTO		
Conteúdo	Valor	Comentário
Edição Semidiplomática	Nome/Não	
Revisão da Edição Semidiplomática	Nome/Não	Nome de quem fez a atividade. Manter o padrão no nome
Edição XML	Nome	
1ª revisão da edição XML	Nome/Não	Acrescentar o nome junto ao nome existente. Ex: Igor Leal / Seu nome
Revisão Final da Edição XML	Sim/Não	
POS	Sim/Não	
Revisão POS	Nome/Não	Nome de quem fez a atividade. Manter o padrão no nome
Anotação Sintática	Nome/Não	
Revisão da Anotação Sintática	Nome/Não	
LEGENDA		
ATENÇÃO NA CORREÇÃO CAMPOS NOVOS		

Figura 1 – Proposta para revisão dos metadados

Fonte: CE-DOHS

Houve uma mudança na ordem de conteúdos, ou seja, passam a ter uma mesma ordem em todas as cartas. Para alguns campos, só foram definidas correções, porém outros foram acrescentados, como escolaridade, habilidade de escrita, profissão, dentre outros. Em relação ao trabalho de revisão das cartas, para isso é necessário abri-la no e-Dictor, e para gerar os metadados é da mesma forma. Aberta, clica-se na opção “documento” e na opção “metadados” ou diretamente pelo atalho CTRL+ALT+M. A Figura 2, a seguir, ilustra uma dessas etapas.

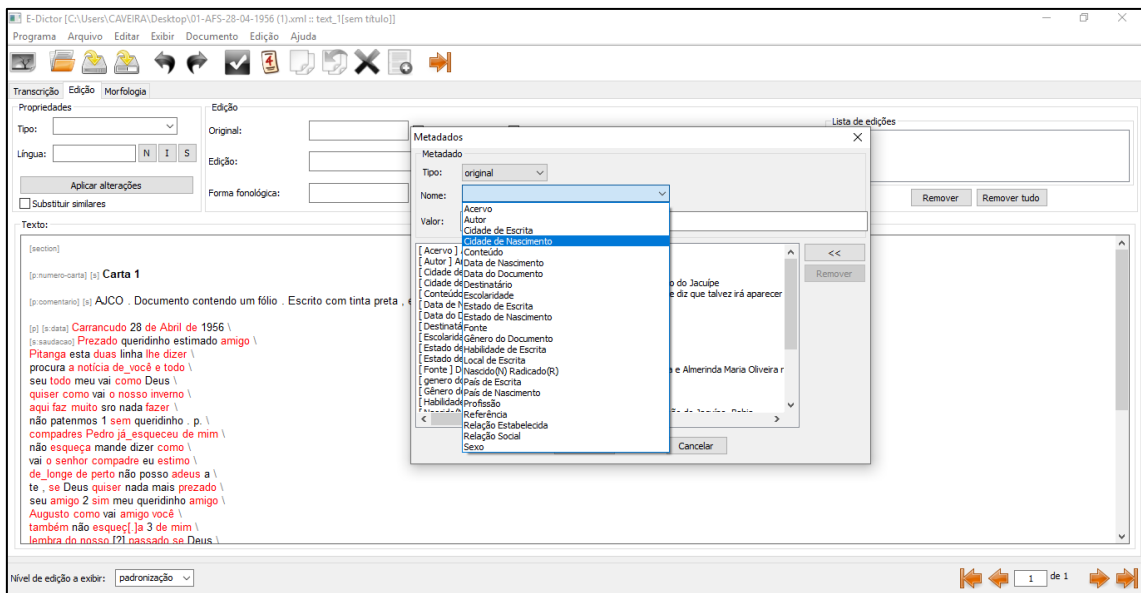


Figura 2 – Campo dos metadados, em processo de revisão, aberto no e-Dictor

Fonte: CE-DOHS

Aberto o campo dos metadados, temos dois tipos: “original” (dados do documento) e “processamento”. Em cada tipo, respectivamente, há uma função, que é obter as informações pessoais do autor e da carta e as atividades realizadas em cada carta e quem as produziram. Depois de aberto o campo dos metadados e aberto o tipo

“original”, existe a opção “Nome”, que são os conteúdos, para que se escolha o campo em que será colocada a informação. Depois de escolhido o campo, a informação será colocada em “Valor”. Para que seja efetuada a alteração, deve-se clicar nas duas setas ao lado direito e para remover algum conteúdo que não esteja na tabela, clica em “Remover” ao lado direito também, depois em “ok” e salva a carta. Na configuração do programa, os campos estão em ordem alfabética, mas é necessário seguir a ordem da tabela de alterações já mencionada. Ao serem colocadas no site *Mãos inábeis*, as fichas catalográficas aparecerão na ordem em que foram geradas no programa e-Dictor.

CONSIDERAÇÕES FINAIS

Os documentos escritos são os materiais necessários para os estudos sócio-históricos da língua e sua disponibilização deve contemplar informações que extrapolam a dimensão interna da escrita. Considerando isso, este trabalho de melhoria na disponibilização das informações sobre *onde*, *quando* e *quem* escreveu os documentos, facilita o trabalho dos futuros pesquisadores, que utilizarão o acervo *Cartas em Sisal*. A atualização das informações, com inserção de novos campos e mudança nos já existentes, foi realizada de forma minuciosa, em todos os documentos que compõe o *corpus*. Considerar as práticas de escrita e as particularidades socioculturais dos redatores ajuda, por exemplo, para a compreensão dos seus processos de letramento, possibilitando investigações em torno de uma reconstituição linguística e sócio-histórica mais aproximada da língua.

REFERÊNCIAS

- BANZA, Ana Paula; GONÇALVES, Maria Filomena. *Roteiro de História da Língua Portuguesa*. Évora: UNESCO Chair in Intangible Heritage and Traditional Know-How: Linking Heritage. University of Évora, 2018. Disponível em: https://dspace.uevora.pt/rdpc/bitstream/10174/22196/1/Roteiro_de_Histo%CC%81ria_da_Li%CC%81ngua_Portuguesa.pdf. Acesso em: 19 ago. 2020.
- GONÇALVES, Maria Filomena; BANZA, Ana Paula (Org.). Introdução. *Património textual e Humanidades Digitais: da antiga à nova filologia*. Évora: CIDEHUS, 2013.
- MATTOS E SILVA, Rosa Virgínia. Reflexões e questionamentos sobre a constituição de corpora para o projeto Para a história do português brasileiro. In: DUARTE, M. E. L.; CALLOU, D. (Org.). *Para a história do português brasileiro – Notícias de corpora e outros estudos*. v. 4. Rio de Janeiro: Faculdade de Letras da UFRJ; FAPERJ, 2002. p. 17-28.
- PAIXÃO DE SOUSA, Maria Clara; KEPLER, Fabio Natanael. *E-Dictor: Uma ferramenta integrada para a anotação de edição e classe de palavras*. VI Encontro de Linguística de *Corpus*, São Paulo, 2007.
- SANTIAGO, Huda da Silva. *A escrita por mãos inábeis: uma proposta de caracterização*. 2019. 722f. Tese (Doutorado em Língua e Cultura) – Instituto de Letras, Universidade Federal da Bahia, Salvador, 2019.