**UNIVERSIDADE ESTADUAL DE FEIRA DE SANTANA**
Autorizada pelo Decreto Federal nº 77.496 de 27/04/76
Recredenciamento pelo Decreto nº17.228 de 25/11/2016
**PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO**
*COORDENAÇÃO DE INICIAÇÃO CIENTÍFICA*

**XXV SEMINÁRIO DE INICIAÇÃO CIENTÍFICA DA UEFS**
**SEMANA NACIONAL DE CIÊNCIA E TECNOLOGIA - 2021**

# EDIÇÃO XML DA CARTA DE CAMINHA (1500) COM BASE EM JAIME CORTESÃO

**Bruno Lopes de Almeida[1]; Zenaide de Oliveira Novais Carneiro[2]**
1. Bolsista PIBIC/CNPq, Bolsista PIBIC/CNPq, Graduando em Letras com Língua Inglesa, Universidade Estadual de Feira de Santana, e-mail: brunouefs@outlook.com
2 Orientadora, Departamento de Letras e Artes (DLA), Universidade Estadual de Feira de Santana, e-mail: zenaide.novais@gmail.com

**PALAVRAS-CHAVE:** Carta de Caminha; eDictor; Linguística de Corpus

## INTRODUÇÃO

A discussão se divide em seções, enumeradas respectivamente, que discutem sobre as partes fundamentais ao desenvolvimento do plano de trabalho: o Projeto CE-DOHS, a ferramenta de edição eletrônica eDictor, o documento e o processo de edição.

## MATERIAL E MÉTODOS OU METODOLOGIA (ou equivalente)

O material, como descrito na Introdução, é a Carta de Caminha, do século XVI, a partir dos estudos de Jaime Cortesão.

A edição XML será feita usando o eDictor, que se trata de uma ferramenta computacional desenvolvida por Kepler, Paixão de Souza e Faria (2007), para possibilitar a edição eletrônicade textos em seus mais variados níveis linguísticos.

## RESULTADOS E/OU DISCUSSÃO (ou Análise e discussão dos resultados)

## O PROJETO CE-DOHS

O Projeto CE-DOHS: Corpus Eletrônico de Documentos Históricos do Sertão integra o Núcleo de Estudos de Língua Portuguesa (NELP) do Departamento de Letras e Artes (DLA) da Universidade Estadual de Feira de Santana (UEFS) é financiado pela Fundação de Amparo à Pesquisa do Estado da Bahia (FAPESB) e foi fundado no ano de 2012. Trata-se de uma grande base de documentos linguísticos referentes ao Português Brasileiro (PB) que busca, através da conexão entre a filologia tradicional e a computação

moderna, nesse caso, fazendo uso da linguagem XML, por meio do editor eletrônico eDictor (PAIXÃO DE SOUSA; KEPLER; FARIA, 2010), contribuir para os estudos linguísticos, históricos e sociais do Brasil. A plataforma foi desenvolvida por Igor Leal (UNICAMP) e Pablo Faria (UNICAMP).

Sob a coordenação das professoras doutoras Mariana Fagundes de Oliveira Lacerda e Zenaide de Oliveira Novais Carneiro, o CE-DOHS fornece à comunidade dois grupos documentais distintos, que se dividem, especificamente, pelo eixo do tempo. O primeiro grupo é constituído por textos (em sua maioria cartas) de autoria de escreventes brasileiros, situados entre os séculos XIX e XX e amostras de fala oral obtidas nos anos 90, totalizando 1553 documentos devidamente editados e disponíveis (http://www.tycho.iel.unicamp.br/cedohs/corpora.html). O segundo conjunto, que abarca o presente plano de trabalho, pretende fazer um recuo mais profundo na origem e no desenvolvimento do Português Brasileiro, apresentando manuscritos dos séculos XVII e XVIII ainda em etapa de edição. É importante dizer que, atualmente, o CE-DOHS supera seu objetivo inicial de pesquisa (o sertão baiano) e oferece à comunidade, de maneira gratuita, na rede mundial de computadores, documentos de diversas localidades do Brasil.

## O EDICTOR: A NOVA FILOLOGIA

Diante do avanço tecnológica e do êxodo científico para o mundo virtual, fez-se necessário que os modelos antigos de se fazer filologia fossem aprimorados. Nessa perspectiva, o eDictor, desenvolvido por Maria Clara Paixão de Sousa, Fábio Kepler e Pablo Faria (2007; 2010), surge como uma ferramenta filológica, agora digital, capaz de proporcionar a edição eletrônica e modernizada, a partir da linguagem computacional XML, de textos e documentos que, após suas edições, são disponibilizados para análises e estudos linguísticos. Desenvolvido, inicialmente, para atender ao Corpus Histórico do Português Tycho Brahe (UNICAMP), o eDictor atualmente é ferramenta principal de diversos centros de constituição de corpora linguístico. O programa apresenta uma interface que possibilita diferentes formas de edição, modernização de textos (conservando o original e indicando, controladamente, todas as alterações feitas) e anotações de nível fonológico e morfossintático. Há também a possibilidade de processamento da edição fac-similar, para que se possa acompanhar texto e imagem simultaneamente.

## O DOCUMENTO

Considerada o maior relato sobre a navegação que resulta na invasão do território brasileiro, a Carta de Pero Vaz de Caminha, originalmente intitulada *Carta a El-Rei D. Manuel* é um dos manuscritos mais importantes para o que conhecemos, hoje, como história do Brasil. Localizado no Arquivo Nacional da Torre do Tombo, em Portugal, o documento, escrito por Pero Vaz de Caminha (1450-1500), escrivão da frota de Pedro Álvares Cabral, é datado de 22 de abril de 1500. Juntamente com a Relação do Piloto Anônimo e as cartas de Mestre João, a Carta de Caminha integra a tríade dos documentos que registram e relatam a expedição de Pedro Álvares Cabral ao Brasil.

O manuscrito é um testemunho presencial que tinha por objetivo transmitir ao rei de Portugal, D. Manuel I, todas as experiências vividas pela expedição no Brasil, além de descrever sua geografia, possíveis riquezas e seus habitantes. A carta foi escrita pouco tempo depois do advento do Tratado de Tordesilhas (1494), no período de ebulição do desenvolvimento das ciências marítimas e da disputa por territórios ainda não descobertos, portanto, há uma perspectiva de deslumbre de um "novo mundo". O espaço de tempo coberto pelo manuscrito é de nove dias, se encerrando, segundo o próprio escrevente, em 1 de maio de 1500. Dentro da narrativa é possível construir um imaginário, mesmo que deturpado, da parte litorânea do Brasil (especialmente do sul da Bahia) e dos povos originários indígenas que habitavam a região.

## O PROCESSO DE EDIÇÃO

Com a versão semidiplomática da Carta de Pero Vaz de Caminha, organizada por Jaime Cortesão, em mãos, transcrevi, de maneira fidedigna, a edição para o software Microsoft Word. O livro de Jaime Cortesão, com o advento do tempo, perdeu popularidade, então se fez trabalhar com uma versão física, desgastada. A qualidade do material não comprometeu as informações que nele estavam presentes, proporcionando uma transcrição sem maiores perturbações.

A cor vermelha indica que a palavra sofreu modificações. O eDictor possui onze diferentes tipos de etiquetas possíveis para a edição, sento estas: junção, segmentação, sobrescrito, ilegível, rasurado, subescrito, tachado, pontuação, expansão, padronização e modernizado. As etiquetas devem ser utilizadas hierarquicamente na ordem descrita no texto, caso uma palavra precise de mais de uma modificação. As edições só podem atingir o nível gráficos. Expandir abreviaturas, acrescentar letras, padronizar palavras para a

língua portuguesa atual, juntar ou separar termos. Alterações de níveis sintáticos e semânticos jamais devem ser feitas, pois isso comprometerá a originalidade do documento. Ao fim, foi possível realizar a edição completa de todos os fólios do documento, que, em breve, serão disponibilizados no site do Projeto CE-DOHS (http://www5.uefs.br/cedohs/).

## CONSIDERAÇÕES FINAIS (ou Conclusão)

Creio que os resultados foram satisfatórios, mais do que esperava devido às complicações físicas e emocionais causadas pela pandemia. A edição da Carta de Caminha faz parte de uma nova etapa do Projeto CE-DOHS e, para um maior retroceder na diacronia do português brasileiro, se faz fundamental a edição realizada nesse plano, do primeiro documento escrito na terra de Pindorama, hoje conhecida como Brasil.

## REFERÊNCIAS


ARQUIVO NACIONAL DA TORRE DO TOMBO. Carta de Pero Vaz de Caminha. Disponível em: (http://digitarq.arquivos.pt/details?id= 4185836).

CASTRO, I. Para uma história do português clássico. In: DUARTE, I.; LEIRIA, I. (Org.). Actas do Congresso Internacional sobre o Português. Lisboa: Associação Portuguesa de Linguística, 1996. v. 2. p. 135-150.

CORTESÃO, J. A Carta de Pero Vaz de Caminha. Lisboa: Portugália, 1967.

GALVES, C. A língua das caravelas: periodização do português europeu e origem do português brasileiro. Campinas: Pontes, 2007.

LOURENÇO, V. T. A Obra de Jaime Cortesão, A Carta de Pero Vaz de Caminha (1967), À Luz da Historiografia Linguística. Revista Philologus, Rio de Janeiro, n. 75, p. 430-441, set./dez., 2019.

LOURENÇO, V.T. Carta de Caminha: Contato Linguístico no Brasil Quinhentista à Luz da Linguística Ecossistêmica. Dissertação (Mestrado em Linguística) – Programa de Pós-Graduação em Estudos de Linguagem, Universidade Federal Fluminense. Rio de Janeiro, p. 125, 2019.

MACHADO FILHO, A.V.L; MATTOS SILVA, R.V. (Org.). O Português Quinhentista:Estudos Lingüísticos. Salvador: EDUFBA, 2002.

MATTOS E SILVA, R.V. A Carta de Caminha Testemunho Lingüístico de 1500. Feira de Santana: UEFS Editora, 1996.

PAIXÃO DE SOUSA, M. C.; KEPLER, F. N.; FARIA, P. E-dictor: Novas perspectivas na codificação e edição de corpora de textos históricos. In: Anais do VIII Encontro de Linguística de Corpus, realizado na UERJ, 13 a 14 de novembro de 2009. Rio de Janeiro, 2009. p. 69-105.

PROJETO CORPUS ELETRÔNICO DE DOCUMENTOS HISTÓRICOS DO SERTÃO (disponível em http://www.uefs.br/cedohs/), 2011.

RIBEIRO, I. A mudança sintática do PB é mudança em relação a que gramática?. In: CASTILHO, A. (Org.). Para a história do português brasileiro. Primeiras idéias. São Paulo:Humanitas, 1998. v. 1. p. 101- 119.