

Interpreting the illegible and the unintelligible in private correspondence

Mariana Gomes, Ana Rita Guilherme, Leonor Tavares

Abstract (en): Project CARDS (Cartas Desconhecidas – Unknown Letters) is a tool for both linguistic and historical research that will make available an on-line archive of Portuguese letters written between the mid 16th and the end of the 19th century. The source documents are private original letters written by speakers from all social backgrounds (popular ones included) and the main covered topics are family, love, friendship, extortion and political business.

This paper presents a brief synopsis of the project and its edition method and tools, together with a reflection on the challenges researchers face when editing private writings, always incomplete from an informative perspective.

Abstract (it): Il Progetto CARDS (Cartas Desconhecidas – Lettere sconosciute) è uno strumento per la ricerca linguistica e storica, che vuole rendere disponibile online un archivio di lettere portoghesi scritte tra la metà del XVI e l'inizio del XX secolo. Le fonti sono lettere originali private, scritte da parlanti di diversa provenienza sociale (anche popolare); i temi più affrontati riguardano la famiglia, l'amore, l'amicizia, l'estorsione e gli affari politici.

Questo contributo presenta una breve sinopsi del progetto, del metodo e degli strumenti impiegati nell'edizione, insieme a una riflessione su la sfida a che gli studiosi devono affrontarsi nel processo di edizione, sempre incompiuto da una prospettiva informativa.

A interpretação do ilegível e do ininteligível na edição da correspondência privada

Ana Rita Guilherme, Mariana Gomes, Leonor Tavares

1. CARDS – Cartas Desconhecidas: O Projecto¹

O projecto CARDS pretende disponibilizar *on-line* manuscritos portugueses contendo cartas privadas escritas entre meados do século XVI e finais do século XIX. Os documentos encontram-se inseridos em processos-crime dos séculos XVIII e XIX do tribunal supremo do reino (a *Casa da Suplicação*²) e em processos dos séculos XVI a XVIII do tribunal católico que julgava e punia heresias (o tribunal do *Santo Ofício da Inquisição*). Enquanto provas judiciais, serviram no seu tempo para ilibar ou condenar réus, formando hoje um corpus documental de ampla representatividade cultural e linguística. Acontece que os autores da epistolografia em causa provinham de ambientes sociais muito diversos, além de que o conteúdo textual de cada carta também era muito variável, cobrindo tópicos que iam da família à amizade, da extorsão à política.

1 Projecto financiado pela Fundação para a Ciência e a Tecnologia com a referência FCT/PTDC/LIN/64472/2006. Equipa actual: Rita Marquilhas (coordenação), Rosario Álvarez Blanco e Ana Maria Martins (peritagem linguística), José Vicente Serrão e Fátima Sá (peritagem historiográfica). Colaboradores: Cristina Albino (paleografia), Edward Vanhoutte e Ron Van den Branden (informática para a filologia). Bolseiras: Ana Guilherme, Leonor Tavares e Mariana Gomes.

2 Só parecem ter sobrevivido em arquivo os processos-crime de finais do século XVIII e os do século XIX anteriores a 1834, data em que o tribunal foi extinto na sequência de uma reforma administrativa (v. Azevedo 1915).

O método que o CARDS utiliza é o da edição electrónica acompanhada de disponibilização do *facsimile* das cartas e de opções de pesquisa on-line. Para tal foi concebida uma base de dados em formato XML seguindo as convenções de codificação concebidas pelo consórcio TEI (Text Encoding Initiative³) e pelo projecto DALF (Digital Archive of Letters in Flanders⁴), projecto este que serviu de inspiração ao CARDS por também fazer edições epistolográficas.

Os objectivos do projecto CARDS passam por tratar tais manuscritos enquanto fontes tanto históricas como linguísticas e, em termos mais imediatos, por conceber uma ferramenta de pesquisa que permita explorar em diversos sentidos os dados oferecidos pelo texto das cartas.

Sinteticamente, o processo de edição passa pelas seguintes etapas:

- i. Edição electrónica do manuscrito num ficheiro XML independente, com seu número de código privativo, contendo informação extra-textual diversa, transcrição semi-diplomática da carta em causa e, em alguns casos, edição modernizada e tradução para uma língua estrangeira;
- ii. Criação de um ficheiro de imagem com o *facsimile* da carta, o qual fica associado ao documento XML;
- iii. Criação de uma base de dados biográfica, também em XML (CDD – Cards Demographic Database) que contém informação de carácter social e biográfico sobre os indivíduos que participaram no acto comunicativo de cada carta.

1.1 A Edição electrónica

As razões pelas quais se optou pelo formato XML e pelas convenções do consórcio TEI são as seguintes:

- i. A linguagem de marcação XML proporciona grandes vantagens ao nível de processamento e de marcação etiquetada de texto. O XML cria ficheiros hierarquicamente organizados que são legíveis quer pela máquina quer pelo homem, servindo assim para vários fins. Além disso, este formato obedece a regras gerais de boa-formação (DTD-Document Type Definition) que são definidas pelo próprio autor da edição que estiver a ser feita, o qual também define com total independência a transformação do texto etiquetado em texto editado, podendo construir a sua própria folha de estilo;
- ii. TEI significa *Text Encoding Initiative* e é um consórcio que tem como objectivo desenvolver e manter um padrão de representação de textos para edições digitais. O TEI desenvolveu linhas de orientação na codificação de textos que se orientam especificamente para o âmbito académico das Ciências Humanas, pelo que é de bastante utilidade nos projectos que queiram manter um elevado padrão filológico no desenvolvimento das suas investigações.

1.1.1. Como funciona a linguagem XML

Esta é uma linguagem em que se atribuem etiquetas a porções de texto, ficando o editor livre de decidir qual o conteúdo das referidas etiquetas. Permitem que o texto etiquetado seja classificado consoante a categoria a que pertence (*elemento*), dentro de hierarquias definidas pelo editor, e consoante as qualidades que exhibe (*atributo*), também dentro de um universo de referência fixado pelo editor.

3 <http://www.tei-c.org/index.xml>.

4 <http://www.kantl.be/ctb/project/dalf/dalfdoc/introduction.html>.

Por exemplo, o elemento <seg> nas convenções TEI remete para um segmento textual independente. Na construção de um ficheiro XML para uma carta do projecto CARDS, usa-se a etiqueta <seg>...</seg> juntamente como os atributos *type*="formulaic" e *n*="protocol/eschatocol", marcando-se deste modo a parte do discurso da carta que obedece a formulários epistolográficos:

```
<seg type="formulaic" n="protocol">Muito festejei ouvir novas vossas</seg>
```

↙ ... ↘

```
<seg type="formulaic" n="eschatocol">Com esta não vos enfado mais</seg>
```

1.1.2. Um ficheiro XML – partes principais

a) O Cabeçalho <teiHeader> – O cabeçalho contém toda a informação extra-textual sobre o discurso da carta. Os elementos do cabeçalho referem as pessoas e as instituições envolvidas no projecto editorial, as regras de transcrição que seguimos, um breve comentário à carta numa perspectiva social, histórica e linguística, detalhes sobre o contexto histórico em que carta foi escrita, bem como a descrição física do manuscrito.

b) O Texto <text>

A parte texto contém a transcrição semi-diplomática da carta disponibilizada pelo projecto CARDS e, em alguns casos, apresenta também a edição modernizada, bem como traduções.

1.1.3. A transcrição

(i) transcrevem-se com etiquetas privativas os elementos textuais do sobrescrito (nome, morada, por vezes também saudação);

(ii) transcreve-se o corpo da carta, o qual está dividido em três partes:

- a) *opener*: contém os elementos de abertura, tais como morada, data e saudação;
- b) texto central da carta;
- c) *closer*: tem os elementos de fecho, tais como morada, data, saudação e assinatura.

(iii) **texto central da carta: o que é que codificado?**

Codificam-se as partes de texto convencional (protocolo e escatocolo), emendas e cancelamentos

1.2. Disponibilização do corpus *on-line*

O sítio do CARDS irá estar ligado ao do Centro de Linguística da Universidade de Lisboa, se bem que por ora, na fase experimental em que está, esteja alojado num endereço provisório⁵. Eis um breve resumo das possibilidades que o sítio da Internet do projecto CARDS oferece:

- a) Ferramenta de pesquisa: este instrumento garante uma pesquisa de dados livre, independentemente da área do conhecimento de onde provém o utilizador; pode pesquisar-se por palavras-chave (que foram identificadas por especialistas em Linguística e em História), por ano, por termos que estejam no corpo de texto e pelo nome de cada participante.
- b) *Downloads*: os ficheiros disponíveis para serem descarregados são os que tornam possível a construção e transformação de cada ficheiro xml: o CARDS.XSL e o DALF.DTD. Desta maneira, é impossível que quem esteja a trabalhar em xml tenha diferentes versões destes ficheiros.
- c) Administração: esta é uma área restrita que permite sobretudo adicionar e remover cartas editadas e seu facsimile; permite também acrescentar entradas à CARDS Demographic Database, disponibilizar versões novas da folha de estilo e actualizar informação dispersa pelos separadores do sítio (bibliografia, monografias e sítios relacionados).

2- Editar cartas interpretando os seus *não ditos*

O trabalho de investigação desempenhado por quem mais directamente trabalha com os documentos deste projecto passa por várias fases. Um primeiro momento é o da localização dos manuscritos, localização essa que implica uma prolongada busca em processos judiciais que se amontoam aos milhares nos arquivos da Casa da Suplicação e do Tribunal do Santo Ofício, ambos depositados no Instituto dos Arquivos Nacionais-Torre do Tombo (Lisboa). Nesta fase torna-se necessário detectar papéis que contenham textos do foro privado no meio de muita documentação burocrática. Após a localização das cartas os investigadores começam a respectiva transcrição, tarefa essa que implica decifração e codificação. Terminado este segundo momento de tratamento da carta, é necessário proceder à leitura do processo judicial que ela integra a fim de melhor compreender aquele evento epistolográfico antigo. Uma vez transcrita e contextualizada, cada carta pode então converter-se, aos olhos do investigador, numa peça que o ajudará na sua reflexão mais ampla sobre as fontes escritas e o seu significado histórico-cultural.

Mas nada é linear no processo de investigação acima esquematizado. Acontece que para decifrar, transcrever, contextualizar e interpretar uma fonte primária é necessário recorrer aos instrumentos teóricos desenvolvidos pela *crítica textual*, pela *história cultural* e pela *análise do discurso*. Um bom exemplo desta necessidade interdisciplinar encontra-se na abordagem do tema que serve de elemento agregador dos artigos deste número da revista *Mnemosyne*, i.e., o *não dito* nos textos privados.

2.1. O cruzamento interdisciplinar e a conversão de fragmentos escritos em textos inteligíveis

Em termos gerais, pode dizer-se que tanto a *crítica textual*, como a *história cultural*, como ainda a *análise do discurso* ajudam o estudioso que lida com fontes escritas a construir várias séries de *não ditos*, implícitos nesses mesmos textos. A crítica textual, na sua acepção mais tradicional, permite edições em que se consiga “manter a palavra original do autor” (Blecuca 1983:10) sem comprometer totalmente a legibilidade do documento junto do público de hoje (Contini 1986:9). Os seus métodos são extremamente valiosos em trabalhos que pretendam olhar para os textos escritos antigos enquanto dados para o estudo da mudança linguística, estudo esse que exige materiais linguísticos genuínos, não mistificados, sob pena de se estar a construir hipóteses equivocadas sobre as gramáticas e os léxicos do passado. O mesmo vale para os estudos

⁵ <http://www.salavisa.net/cards/index.php>

históricos e discursivos que foquem a estreita relação entre propriedades materiais e textuais dos textos escritos e o ambiente em que se moviam os respectivos autores e leitores.

No projecto CARDS, e usando os métodos da crítica textual, enfrentam-se universos implícitos, *não ditos*, como os seguintes:

- o *conhecimento linguístico* do falante que escrevia o enunciado de cada carta (assumindo uma perspectiva formal da língua do autor da carta, na linha da investigação Chomskiana, cf. Chomsky 1986);
- a *cultura* comum aos elementos da comunidade em que cada carta circulava (assumindo uma perspectiva histórico-cultural da escrita, na linha de estudos como os de Roger Chartier, cf. Chartier 1992/1997);
- o *contexto de uso* construído pelo discurso das cartas, sobretudo por aquelas que se trocavam em cadeia (assumindo a perspectiva social da língua de uma comunidade tal como ela é vista pelos investigadores da análise do discurso, cf. Van Dijk 1997).

a) A construção do conhecimento linguístico

Tendo em vista o conhecimento linguístico dos falantes de épocas passadas, procede-se no projecto CARDS à distinção entre edições semi-diplomáticas e edições interpretativas (Roncaglia 1975). O primeiro tipo de edição permite que se capte nestas fontes epistolográficas originais, por exemplo, a forma como certos autores menos cultos se distanciavam de ortografias adoptadas nos textos impressos, testemunhando assim traços fonológicos e morfológicos, bem como o léxico de variedades não-standard do português. Já a edição interpretativa dos mesmos documentos, com sua razoável normalização ortográfica, permite preparar os textos para serem tratados mecanicamente enquanto *corpus* histórico-dialectal, pesquisável em termos sintáctico-semânticas. É verdade que, por ora, os textos do CARDS não formam um *corpus* etiquetado no sentido em que o são outros *corpora* do português histórico, como o Corpus Informatizado do Português Medieval⁶ ou como o Tycho Brahe⁷, mas cada carta está a ser indexada com um conjunto de palavras-chave que permitem que ela venha a ser útil a investigadores que busquem fenómenos específicos de mudança linguística. Além disso, a ficha biográfica que se elabora para cada autor de carta permite ter uma ideia razoável de como se caracterizavam sociologicamente aqueles escribas, caracterização essa que se torna pertinente no momento de interpretar o seu comportamento linguístico - um sêssio, um betacismo, uma concordância inesperada ou uma ordem de palavras intrigante.

b) A construção da cultura

Adopta-se neste projecto um conceito de *cultura* bastante amplo, como o que propôs Roger Chartier: “a totalidade das linguagens e das acções simbólicas que são próprias” de uma comunidade, veiculando crenças, valores e juízos de valor sobre si e sobre os outros (Chartier 2004:14). Esta abordagem tira um partido máximo da reprodução facsimilada dos documentos porque se admite que isso permite ao estudioso, de certa forma, “ouvir os mortos com os olhos” (Quevedo, *apud* Chartier 2007). O efeito de sentido que cada autor de um texto escrito pretende com o seu enunciado cabe sempre, também, na materialidade do documento, a qual vai potenciar ou limitar significados. Por outro lado, as competências culturais do leitor que recebe o texto são um novo entrave aos sentidos que o autor queria produzir. De modo que o historiador cultural, se quer chegar às crenças de uma comunidade distante no tempo tem de ultrapassar, também ele, a leitura estritamente textual das suas fontes.

No caso dos documentos editados neste projecto, são sobretudo pesquisáveis fenómenos culturais como a expressão da delicadeza e da identidade, ou as táticas para a superação da distância; adicionalmente, e por serem cartas integrando processos judiciais, estão nelas bem documentados os fenómenos da criminalidade e da culpa. Cruzando textos coevos e materialidades do suporte, já foi possível produzir algumas monografias nesta linha, centradas precisamente nos documentos que o projecto agora edita⁸.

6 <http://cipm.fcsh.unl.pt/login.jsp>.

7 <http://www.tycho.iel.unicamp.br/~tycho/corpus/index.html>.

8 V., por exemplo, Marquilhas 2008.

c) A construção do contexto de uso

Uma terceira modalidade de estudo do *não dito* em cartas privadas do passado consiste em contextualizar microscopicamente o momento comunicativo em que foram produzidas. Os tribunais que guardavam as cartas enquanto provas de um crime também interrogavam os réus que as tinham escrito ou recebido. As actas desses interrogatórios (*autos*) figuram nos mesmos processos judiciais que contêm provas (como são as cartas), de modo que toda uma possibilidade se abre de fazer estudos de análise histórica do discurso à custa desta rica documentação. Partindo de um conceito de discurso enquanto “uso da linguagem, socialmente situado, interactivo, e em que se comunicam crenças” (Van Dijk 1997: 2) será possível investir na relação entre implícitos conversacionais (Grice 1975) e enunciados efectivamente produzidos para observar como os falantes construíam, em diferentes épocas, a noção de contexto de uso da língua. Ainda nenhuma monografia foi produzida pelo CARDS nesta perspectiva de análise, mas a sua possibilidade está aberta porque os investigadores do projecto fazem sistematicamente um resumo de cada processo judicial com que trabalhem e que contenha cartas.

Em suma, e apesar da sumária apresentação, julgamos ter deixado claro que o projecto CARDS pretende começar a constituir um *corpus* de Português *não standard* a partir de textos informais feitos em várias épocas da história recente da língua. Apura-se assim um instrumento que servirá em seguida para formular hipóteses sobre mudanças na língua e na cultura portuguesa ao longo da Idade Moderna e no início da Contemporânea.

Referências :

- Pedro de AZEVEDO, *Ano e meio de vida do arquivo dos Feitos Findos*. Separata dos *Anais das bibliotecas e arquivos de Portugal*. Coimbra: Imprensa da Universidade, 1915.
- Alberto BLECUA, *Manual de crítica textual*. Madrid: Editorial Castalia, 1983.
- Roger CHARTIER, “A ‘Nova’ História Cultural Existe?”, in *Cultura: revista de história e teoria das ideias*, 2004, 9-22.
- Roger CHARTIER, *A ordem dos livros*. Lisboa: Editora Passagens, 1992/1997.
- Roger CHARTIER, “L’écrit et l’écran, une révolution en marche” ✓ Conferência proferida no Collège de France na inauguração do ano lectivo 2007/2008\, in *Le Monde*, 12.10.2007.
- Noam CHOMSKY, *Knowledge of Language*. Nova Iorque: Praeger, 1986.
- Gianfranco CONTINI, *Breviario di Ecdotica*. Milão: Riccardo Ricciardi Editore, 1986.
- Paul GRICE, “Logic and conversation”, in P. Cole e J. L. Morgan (eds.), *Syntax and Semantics*, 3. Nova Iorque: Academic Press, 1975, 41-58.
- Rita MARQUILHAS, “‘Eu ainda sou vivo’. Sobre a edição e análise linguística de cartas de gente vulgar”, *Estudos de Linguística Galega*, 1, 2008.
- Aurelio RONCAGLIA, *Principi e Applicazioni di Critica Testuale*. Roma: Bulzoni Editore, 1975.
- Teun A. VAN DIJK, “The study of discourse”, in Teun A. van Dijk (ed.), *Discourse as Structure and Process*. Londres: Sage, 1997, 1-34.