

## DICIONÁRIO, ARQUIVO E DOCUMENTAÇÃO

Maria Teresa MARTINS<sup>1</sup>  
mate\_martins@yahoo.com.br  
Universidade Estadual Paulista (UNESP)  
Sessão de comunicação 1: Língua, Hiperlíngua e Arquivo  
Coordenadores: Marisa Grigoletto (USP) e Rosely Diniz Machado (FURG)

### Introdução

A partir da perspectiva teórica da Análise de Discurso, em articulação com a História das Idéias Lingüísticas, os dicionários são tomados como objetos histórico-discursivos cuja produção é determinada pela conjuntura sócio-histórico-ideológica em que se inserem.

Baseados nessa noção de dicionário e na noção de *arquivo*, como “espaço polêmico das maneiras de ler” Pêcheux (1997, p.57), é que trabalhamos na informatização por imagem do *Dicionário Brasileiro da Língua Portuguesa*, de Macedo Soares (1888), no âmbito do Projeto “Documentação lingüística: arquivo, instrumentação e divulgação” (FAPESP, 2006/00234-8), coordenado pelo Prof. Dr. José Horta Nunes. O objetivo geral desse projeto é efetuar uma reflexão sobre as noções de *documentação lingüística* e de *arquivo*, a partir da perspectiva da Análise de Discurso e da História das Idéias Lingüísticas. Entre os objetivos específicos, consta a instrumentação de alguns dicionários de língua portuguesa. De acordo com Nunes (2006),

A instrumentação, além de implicar a produção de uma nova versão, afeta também a circulação do texto. Quando impressos, os dicionários são de acesso restrito, freqüentemente alocados em bibliotecas de obras raras (o que exige pesquisa *in loco*), onde muitas vezes não são permitidas cópias xerográficas. As novas versões eletrônicas condicionam uma circulação mais ampla.

Nosso objetivo hoje é apresentar o trabalho de instrumentação de Soares (1954/1955). Primeiramente, apresentaremos o dicionário que foi objeto da pesquisa. Em seguida, discutiremos a questão da informatização.

### ***Dicionário Brasileiro da Língua Portuguesa, de Macedo Soares (1888)***

Nunes (2002) apresenta um estudo sobre a dicionarização no Brasil. De acordo com o autor, o processo de dicionarização no Brasil pode ser resumido em quatro momentos: 1. transcrição e

---

<sup>1</sup> Mestre em Estudos Lingüísticos pela Unesp – São José do Rio Preto.

comentário de termos (em que se encontram os relatos de viajantes); 2. listas temáticas de palavras; 3. dicionários bilíngües português-tupi (Época Colonial – XVI-XVII); 4. dicionários monolíngües (os portugueses que foram mais tarde *transferidos* para território brasileiro (Moraes (1789), Aulete (1881), Figueiredo (1899), etc) e os produzidos no Brasil, entre eles os de complemento (Costa Rubim, 1853), os de regionalismo (Coruja, 1856) e os de brasileirismos (**Macedo Soares, 1888**; Beaurepaire Rohan, 1889)).

O dicionário de Macedo Soares (1888), ao lado do de Beaurepaire Rohan (1889), reuniu um conjunto de brasileirismos que se encontravam até então dispersos em listas anônimas e na literatura de românticos como José de Alencar (*Diva*, 1865). Como observa Nunes (2001, p.82), apresenta uma escrita singular e marca “um momento importante da lexicografia brasileira, que coincide com os movimentos republicanos”. E o autor continua:

No último quarto do século XIX, a política da língua toma um outro rumo. Começa o interesse pelos falares populares, pelo “povo brasileiro”. O povo, ignorado durante a Monarquia, faz irrupção então nos dicionários. Além disso, as fontes escritas integram jornais, leis, revistas, documentos históricos que tocam o cotidiano das cidades. Aparece, assim, um discurso que leva em conta a vida cotidiana. A nomenclatura, além de termos da fauna e da flora, compreende igualmente os termos pertencentes aos domínios da atividade social e econômica tanto no meio rural como no meio urbano. (p.82)

Há, portanto, nesse momento (República), um movimento de reconfiguração de arquivos com relação ao Império: passa-se a representar o povo por meio dos brasileirismos. Temos então que uma mudança no momento histórico leva a uma mudança no modo de ler o arquivo.

### **A informatização**

O dicionário de Macedo Soares foi publicado em 1888 apenas até a letra “C” nos Anais da Biblioteca Nacional (Rio de Janeiro). Apenas em 1954 (Tomo1) e 1955 (Tomo 2) é que o dicionário é publicado na íntegra, graças ao trabalho de Julião Rangel de Macedo Soares que *coligiu, reviu e completou* a obra do pai.

Trabalhamos com essa edição do dicionário de Macedo Soares de 1954/1955, editada pelo Instituto Nacional do Livro, do Rio de Janeiro. Esta escolha marca um primeiro gesto de leitura do arquivo: a seleção da versão. Tendo em vista que nosso objetivo é fazer com que seja possível

pesquisar palavras para conhecer aspectos da época, isso seria impossível com a edição que vai até a letra “C”. Palavras como “escravo”, por exemplo, ficariam de fora do *corpus*.

O dicionário todo foi *scaneado* apenas por imagem<sup>2</sup> utilizando equipamentos (microcomputador e scanner HP) adquiridos pelo projeto acima referido. As imagens não poderão ser divulgadas na internet devido aos direitos autorais, portanto comporão o acervo do projeto estando disponível para pesquisa *in loco*. Ainda assim, pode-se dizer que a informatização altera a circulação da obra, tendo em vista que enquanto obra impressa sua circulação fica restrita, uma vez informatizada o acesso a ela por pesquisadores e leitores interessados será mais viável, apesar do silêncio da/na internet.

Ressalte-se que consideramos que com esse trabalho de informatização não temos a *obra original no computador*, mas sim uma sua *versão* (ORLANDI, 2001, p.13). Considerando que o *texto original* “é uma função da historicidade” e que as alterações na materialidade do texto “correspondem a diferentes gestos de interpretação” (ORLANDI, 1996, p.14).

O que temos ao final, então, é uma passagem do dicionário impresso (“*dictionnaire papier*”) para o *dicionário informatizado* (“*dictionnaire informatisé*”), nos termos de Pruvost (2000, p. 106-107).

Há ainda que se observar que da perspectiva teórica em que nos situamos, questionamos a transparência dos meios. Segundo Wanderley (2003, p.136), “as novas tecnologias chamadas tecnologias da informação são produto de línguas imaginárias (linguagem de programação) funcionando em materialidades de silício, de energia, e em rede”. Nossa tarefa é não apagar a materialidade da língua.

Uma maneira de fazê-lo é desautomatizar leituras. À medida que esse dicionário, e outros que são informatizados pelo projeto (Costa Rubim (1853), Coruja (1856) e Beaurepaire Rohan (1889)), ofereçam possibilidades de buscas, cada leitor construirá seu percurso de leitura pelo texto. Segundo Orlandi (2003, p.14), “Pela tomada em consideração da materialidade do texto é que se

---

<sup>2</sup> Tendo em vista um dos objetivos do projeto, qual seja, de informatizar “por imagem e texto” os dicionários de Coruja (1852) e de Rubim (1953) e “somente por imagem” os dicionários de Soares (1888) e de Rohan (1889).

apresenta a variação, em outras palavras, as **versões** de leitura possíveis. **A leitura é aferição de uma textualidade no meio de outras possíveis.** Há textos possíveis às margens do texto” (grifo nosso).

### Observações finais

“Os tesouros de ontem ressuscitam nas tecnologias de amanhã.” (PRUVOST, 2000, p. 129).

Embora não vejamos os dicionários como tesouros da língua, mas como instrumentos lingüísticos e como objetos histórico-discursivos, acreditamos que as tecnologias de leitura-escritura de arquivo que a contemporaneidade nos coloca fazem com que haja a possibilidade de circulação, de leituras, de versões, de re-inscrição na memória, de análises.

No caso específico do dicionário de Soares, o que marca uma outra versão é i. o meio de circulação; ii. as mudanças materiais que são também gestos de leitura, como limpar sujeiras das páginas que possam dificultar a leitura e, por outro lado, manter carimbos e outras marcas que deixam visíveis os percursos por onde o dicionário passou.

### Referências

- NUNES, J.H. Dicionarização no Brasil: condições e processos. In: Nunes, J.H.; Petter, M. (org.) *História do saber lexical e constituição de um léxico brasileiro*. São Paulo: Humanitas/ FFLCH/ USP: Pontes, 2002.
- NUNES, J. H. Léxico e língua nacional: apontamentos sobre a história da lexicografia no Brasil. In: ORLANDI (Org) *História das idéias lingüísticas: construção do saber metalingüístico e constituição da língua nacional*. Campinas: Pontes; Cáceres: Unemat, 2001.
- NUNES, J.H. O texto na documentação lingüística. *Trabalhos em Análise de Discurso*. N.3. São José do Rio Preto : Unesp ; Fapesp, 2006.
- ORLANDI, E.P. *Discurso e texto*. Campinas : Pontes, 2001.
- ORLANDI, E.P. *Interpretação ; autoria, leitura e efeitos do trabalho simbólico*. Petrópolis : Vozes : 1996.
- ORLANDI, E.P. Ler a cidade: o arquivo e a memória. In: Orlandi, E.P. (org.) *Para uma enciclopédia da cidade*. Campinas: Pontes, 2003.
- PECHEUX, M. “Ler o Arquivo Hoje”. In: Orlandi, E.P. (org.) *Gestos de leitura: da história no discurso*. 2.ed. Campinas: Editora da Unicamp, 1997.
- PRUVOST, J. *Dictionnaires et nouvelles technologies*. Paris: Presses Universitaires de France, 2000.
- SOARES, A. J. M. *Dicionário Brasileiro da Língua Portuguesa*. Rio de Janeiro: Instituto Nacional do Livro, 1954.
- WANDERLEY, C. Recursos do discurso eletrônico e a construção do sistema Endici. In: Orlandi, E.P. (org.) *Para uma enciclopédia da cidade*. Campinas: Pontes, 2003.