

## REFLEXÕES SOBRE MÉTODOS E ADEQUAÇÕES TEÓRICAS PARA TRATAMENTO DO LÉXICO EM PERSPECTIVA HISTÓRICA

Américo Venâncio Lopes MACHADO FILHO  
Universidade Federal da Bahia/Grupo PROHPOR

### PREAMBULANDO

Para Rosa Virgínia Mattos e Silva (Mattos e Silva, 2006, p.17), é o período arcaico do português um momento histórico da língua em que “ainda não se explicitara a norma, os padrões de uso prestigiado, estabelecidos pelos gramáticos”. Como se sabe, a nova tendência que progressivamente seria engendrada pelas sociedades neolatinas, que passariam a reconhecer no romance um quê de prestígio linguístico em face do latim, só viria a se manifestar mais explicitamente, a partir do século XVI e de forma ainda incipiente, se comparada aos padrões normalizadores modernos.

Esse dado corrobora com a noção de ser o trabalho lexicográfico histórico um campo de pesquisa deveras idiossincrático, já que, diferentemente da lexicografia contemporânea, a conservação da diversidade de usos da escrita, isto é, o pleno registro da variação gráfica, é muito mais requerido naquela do que nesta, passando a se configurar como uma das balizas metodológicas norteadoras do trabalho de pesquisa diacrônico do léxico, nomeadamente no que se refere à construção de dicionários históricos da língua, sobretudo aqueles que objetivam registrar o período que antecede as novas posturas sociais e linguísticas dos anos quinhentos.

Em detrimento ao difundido princípio linguístico hodierno de seleção baseado em frequências de uso, o processo de lematização de unidades lexicais deve, no trabalho de viés histórico, por exemplo, desviar-se dos preceitos de canonização dos signos lematizados, com que lidam os lexicógrafos contemporâneos.

Na lexicografia histórica a conformação dicionarística dos lemas deve ganhar contornos, não exclusivamente pela sua “face neutra”, isto é, não apenas pela forma flexionalmente vazia do lexema, como é hoje feito, mas pela variedade das formas, quer simples, quer compostas ou complexas, ainda textuais, que possam ocorrer nos *corpora* sem um correspondente morfológico canônico. Isso vale dizer que se um item lexical só ocorre uma única vez no feminino plural, o lema deverá corresponder a essa mesma forma, em prol da manutenção do real espólio linguístico da época que se investiga, sem qualquer prejuízo para o método. São uma exceção, obviamente, os verbos, que, por exibirem um

comportamento flexional bastante prolífico e produtivo, podem e devem conformar-se aos ditames tradicionais de lematização.

O desenho da macroestrutura de um dicionário histórico do português, notadamente de seu período arcaico, deve, então, privilegiar, para além desse procedimento antes sugerido de lematização, um sistema de remissão, de alguma forma perdulário, que possa arcar com grande parte da exuberância gráfica existente, sem deixar de permitir ao público-alvo uma consulta rápida e eficaz às unidades léxicas de seu interesse. Essa estratégia possibilitaria, ainda, que o provável desconhecimento, por parte do consulente, da forma ou das formas gráficas de lexias de um período distante, não lhe obliterasse uma resposta adequada do dicionário à sua curiosidade, mesmo quando de alguma maneira pudessem essas formas linguísticas ter sido alteradas substancialmente com o tempo, a ponto de não mais serem identificadas no presente, como é o caso do verbo *ser* < *seer* < (lat. *sedere*) ou da palavra *çapato* (talvez do turco *c(apata)* ~ *sapato*, para servir aqui de exemplo, numa consulta alfabética.

Assim, a nomenclatura deveria idealmente comportar não apenas toda a variação detectada nos *corpora*, mas, também, fomentar uma estratégia de “falsas entradas” em português moderno – somente quando estritamente necessárias – contudo devidamente sinalizadas com indicadores estruturais, tipográficos e não-tipográficos, como elementos facilitadores de consulta.

No tocante à definição, que, para a grande maioria dos metalexígrafos – a exemplo de Guilbert (1969, p. 29) –, é o elemento primordial, basilar e indissociável de qualquer dicionário, embora fosse desejável que se pudesse obedecer ao que prega a lexicografia moderna, para que, entre os vários tipos de definição, a lógica se apresentaria como a idealizada (cf. Biderman, 1993, p. 29) – isto é, aquela que, com base na lexicologia estrutural, se compusesse a partir da equação incontestada do *genus proximum* e das *differentiae* de todo o contínuo de oposição do conteúdo lexical –, a distância temporal que se interpõe entre o linguista histórico e o léxico pulveriza, de certa forma, o quadro sêmico que se poderia construir em sua mais plausível completude, assim como inviabiliza, por vezes, até a codificação da informação semântica numa definição de compromisso, chamada de lexicográfica, em que figurariam apenas o *genus* e as *differentiae* caracterizadoras de cada unidade.

Considerando o que afirma Greimas (1966, p. 36) que “a comunicação é um ato e, por isso mesmo, é sobretudo escolha” e que no “interior de um universo significante a partir do

qual opera, escolhe cada vez certas significações e exclui outras”<sup>1</sup>, recuperar a organização do conteúdo lexical total de um dado item no uso sociolingüístico, em um momento específico da história, traduz-se como improvável, “se se considerar o nível de imprecisão a que se poderia chegar em relação ao próprio nível de conhecimento fragmentário que se tem da sincronia que se pretende (...) caracterizar” (Machado Filho, 2003, p. 21). Por isso, se pode optar por recorrer, quando necessário e quiçá sem remorso, à estratégia da – com razão, tão combatida pela lexicografia moderna, porém por esta utilizadíssima – definição sinonímica, a partir da observação das acepções contextuais, valendo-se, todavia, de paráfrases lexicográficas, quando possível. Caberia aos dicionários modernos, então, nesse caso, providenciar os subsídios sêmicos complementares ou correspondentes para a satisfação da informação semântica desejada por seu leitor.

Ademais, por se estar aqui a tratar de questões de métodos a serem adotados na elaboração de um dicionário histórico, especialmente se este contempla a etimologia, a estrutura dos sentidos ou acepções adotada deve ser, preferencialmente, linear (*flat structure*), em ordenação eminentemente histórica, em que se obedeça à trajetória semântica do item definido, em face da sua ocorrência temporal nos *corpora*.

## 1 REVENDO O PREÂMBULO COM EXEMPLOS

Embora acreditem alguns teóricos, a exemplo de Barbosa (apud Welker, 2004, p. 107) que a microestrutura de um dicionário de língua deva corresponder a um “programa e a um código de informações aplicáveis a qualquer entrada”, para o desenvolvimento de um dicionário histórico nos moldes do ora proposto esse posicionamento se torna contraproducente, no sentido em que, em diversos momentos, a entrada deve ser composta, unicamente, de um lema com indicação remissiva. Ademais, verbos, assim como outros elementos gramaticais, a exemplo de formas dependentes, não devem ter o mesmo tratamento, já que, em relação aos primeiros, interessa à história da língua identificar e registrar as manifestações flexionais diacrônicas que se salvaguardaram nos textos antigos e, no que concerne às formas dependentes – esvaziadas de conteúdo extralingüístico –, o enunciado definitório se dá por inferência à sua categoria gramatical.

Com base no exposto, a microestrutura básica do verbete obedeceria à conformação do diagrama abaixo:

---

<sup>1</sup> Tradução livre do original: *la communication est un acte, et, de ce fait même, elle est surtout choix (...) A l'intérieur de l'univers signifiant à partir duquell elle opère, elle choisit chaque fois certaines significations et en exclut d'autres.*

<lema/> <lema>(<lema secund./> ~ - <lema secund./>)<clas./> .<clas./><etim./> ( ) <etim./> <fonte etim./> <sup>x</sup> <fonte etim./> (<remis./> → .<remis./><def./> ‘ ’<def./>; <def./> ‘ ’. <def./>) <data/tx/loc/> [ ]<data/tx/loc/><abon./> <abon./>(<data/tx/loc/> [ ] <data/tx/loc/><abon./> (...) x (...) <abon./>).

em que o lema, delocado em 1 cm à esquerda, é, em caso de variação gráfica, seguido dos lemas secundários, indicados pelo sinal do til; pela classificação gramatical abreviada por ponto; pela etimologia entre parênteses, a que se apõe a fonte de pesquisa abreviada e em sobrescrito; pela remissão, quando se justifique, indicada por seta e fechada por ponto; pela definição ou definições entre apóstrofes, separadas por ponto-e-vírgula e fechadas por ponto; pela datação, pela indicação do texto de que se extrai a abonação e localização da página, linha ou coluna no original, entre colchetes; e por cada abonação correspondente, com o item em negrito, fechada por ponto final.

Em decorrência desse posicionamento, os verbos mereceriam no planeamento um estrutura diferenciada dos outros itens lexicais.

Na sequência, exemplificam-se o formato dos verbetes de formas nominais, formas verbais, de formas meramente remissivas e das “falsas entradas” antes referidas, através de alguns fragmentos:

**abade** ~ **abbade** – sm. (< lat. *abba*###*em*)<sup>h</sup> ‘prelado responsável por uma abadia’. [1214/tasl/7] E mado q(ue) o **abade** d’Alcobaza lis de aq(ue)sta dezima q(ue) el ten ou teiuer. [xiii/frac/96v] [T]oda carta q(ue) seya feyta ante alguus e seya y posto seello del rey ou de arçabispo ou de bispo ou d(e) **abade** ou d(e) prior ou d(e) concello ou de pessoa conhecida por testimonho, esta ualla, fora se aquel (contra) que) for feyta lala carta a poder desfaz(er) cu) dereyto. [xiii/frac/145r] E possao o monge q(ue)rellar a sseu **abbade** ou a seu mayor so cuyo poder é. [xiv/flos/13vc1] estes lavravam e gaanhavam e colhiam seu pam e envyavam ende a muy mayor parte a este **abade** que a metesse em prol dos pobres.

.....  
**abbade** → abade.

.....  
**çapatos** – sm. pl. (< origem obscura, talvez do turco *c(apata)*<sup>g</sup> ‘calçado, em geral de sola dura, que cobre o pé, parcial ou completamente.’ [xiv/flos/62rc2] na terra d’ouriente hu os clerigos sō sem pecado casados, forom dous clerigos e viviam per fazer **çapatos** e eram vezinhos huus doutros.

.....  
**espedir-se** ~ **espidir-se** – v. (< lat. *expete(re)*<sup>m</sup> ‘despedir-se’; ‘ir embora’; ‘dispensar’. || INF [xiii/frac/120r] E mādamos que o senhur de que alguu fidalgo **se quiser espedir** no) lhy faça por en outro mal, senõ que lhy demãde seu dereyto se quiser e nõno deoste ne) uilte por en. [xiii/frac/120r] Todo vassallo despoys que **se espidir** de seu senhor e non lhy quiser tornar as armas nen os caualos *que* del ouue, possao (o) senhor retar polhas lorigas. [xiii/frac/119v]: E quando quiser **espedirse** del beygelhy a mão [...] ao senhor de que se espede e digalhy: foan tal caualeyro uos mãda beygar a maa e **espedirse** de uos *per* mi. || IP3 [xiii/frac/119v] E quando quiser **espedirse** del beygelhy a mão [...] ao senhor de que **se espede** e digalhy: foan tal caualeyro uos mãda beygar a maa e **espedirse** de uos *per* mi. [xiv/flos/13vc1] mais quando ha de morrer conhoce sa morte e dize-o a todolos frades e **espede-se** deles e morre e dá sa alma a

deus. || IPP3 [xiv/flos/32rc2] ele pois que se maenfestou e ordi)ou seu testamento e **espediu-se** a seus amigos, e a primeira noyte que veo sayu-lhi a alma do corpo. || IPP6 [xiv/flos/37vc2] e pois aqueles maaos conselheiros virom que rem nõ podiam acabar daquelo por que veerom, **espidirom-se** do sancto bispo e do governador e foram-se muyto amaros e muyto tristes pera sa casa.

.....  
**espidir-se** → espedir-se.  
.....

[**sapatos**] → çapatos.  
.....

Como se pode perceber na análise dos exemplos acima o sistema de abonação dos verbos obedece a uma dinâmica própria que procura apresentar ao consulente todas as possibilidades de flexão verbal detectada no *corpus*, a partir de uma hierarquia racional, em que modo, tempo e pessoa (obviamente apenas as formas detectadas) são indicados pelas abreviaturas correspondentes, após barras verticais.

As chamadas “falsas entradas” remissivas, acima representadas pelo verbete [*sapatos*] → *çapatos*, indicam que, embora a lexia pesquisada não esteja atestada na forma gráfica que se encontra patente entre os colchetes, a sua correspondente histórica estaria devidamente lematizada no dicionário, conquanto em forma morfológica de plural, não-canônica, portanto, em função dos dados, que exemplarmente aqui só teria ocorrido com essa configuração linguística no *corpus*.

Esse posicionamento busca colocar a consulta “ao alcance de um público relativamente vasto e não preparado filologicamente para enfrentar os textos medievais à vista desarmada” como diria Castro (1973, p. 5), evitando que formas como *hymno* (*hino*), *erdeyro* (*herdeiro*), *sagramento* (*sacramento*), entre muitas outras, sejam ignoradas no processo de pesquisa, em função da ordenação alfabética linear.

## CONCLUINDO

Se se considerar que a construção do conhecimento lexicográfico em língua portuguesa só se desenvolve mais verticalmente a partir do século XVI, com o Renascimento e que a lexicografia brasileira contemporânea tem-se baseado – por tradição e até ao presente – normalmente em referências arquetípicas de grandes trabalhos precedentes, como o de Morais Silva, entre outros, há-de se compreender que, embora existam hoje obras como o Aurélio e o Houaiss, a metodologia aplicada à elaboração de dicionários de momentos mais

recuados da língua, sobretudo do período arcaico do português, constitui-se, presentemente, em uma área de conhecimento por se explorar.

E é sobre essa questão que se propôs aqui começar a discutir.

## Resumo

No que se refere à produção lexicográfica mais antiga em língua portuguesa, só restou à ação impiedosa do tempo um pequeno dicionário de verbos *latim/português*, provavelmente escrito nos finais do século XIV ou no início do XV. Não obstante essa tímida e tardia estréia do português em tão instigante cenário dos estudos lingüísticos, a construção do conhecimento lexicográfico vem a se desenvolver verticalmente a partir do século XVI, com o Renascimento, e reconhece nos séculos posteriores grandes produções, que viriam a servir de base para a construção de dicionários modernos, que têm, no entanto, por tradição, até o presente, se baseado normalmente em referências arquetípicas de grandes trabalhos precedentes. Essa tendência condiciona-se, todavia, a um universo deveras restrito, se considerado, comparativamente, ao espólio de outras línguas de cultura de sua família lingüística ou mesmo de línguas derivadas do tronco de que se originou a língua portuguesa. Ademais, o resgate histórico do léxico de momentos pretéritos da língua, nomeadamente do período arcaico do português, tem restado à espera de projetos que objetivem reconstruir o passado lexical, a partir de um instrumental teórico consistente, inovador e coerente e, ao mesmo tempo, compatível com os preceitos da lexicografia moderna. É a lexicografia histórica – malgrado o pouco esforço que se empreendeu na elaboração de glossários e vocabulários – uma área relativamente nova entre as ciências do léxico e, por isso mesmo, demanda que se estabeleçam metodologia e fundamentação teórica que, mesmo que por vezes possam parecer antagônicas às praticadas pela lexicografia moderna, sobretudo na composição de corpora, no aproveitamento dos dados e no seu adequado tratamento, representem um avanço no desenvolvimento dessa área do conhecimento. Considerando-se que as características específicas dos textos antigos, sobretudo dos medievais, têm, de certa forma, obliterado a utilização plena dos preceitos metodológicos e dos recursos de programas informáticos, nomeadamente no que se relaciona com a variação gráfica dos vocábulos, pretende-se com este trabalho apresentar uma reflexão sobre os métodos e adequações teóricas com que se tem trabalhado na composição de um dicionário etimológico do português arcaico.

## Referências

BIDERMAN, Maria Tereza. A definição lexicográfica. *Terminologia*, Porto Alegre, n. 10, 1993. pp. 23-43.

BRUCKER, Charles.(1988). *L'étymologie*. Paris: Presses Universitaires de France.

CASTRO, Maria Helena et al. Normas de transcrição para textos medievais portugueses. *Boletim de Filologia*, Lisboa, n. 12, 1973, p. 417-425.

GREIMAS, A. *Sémantique structurale : recherche de méthode*. Paris : Larousse, 1966.

GUILBERT, Louis. Dictionnaires et linguistique: essai de typologie des dictionnaires monolingues français contemporains. *Langue Française*, 2 : 04-29, 1969.

MACHADO FILHO, Américo. Breve incursão pelo léxico medieval do português: o testemunho de um manuscrito trecentista. *Revista Estudos Lingüísticos e Literários*, 29/30: 15-29, 2003.

MACHADO FILHO, Américo. *O manuscrito da livraria 522: edição diplomática*. Coimbra: Faculdade de Letras da Universidade de Coimbra, 2006. [Trabalho de Pós-Doutoramento inédito].

MATTOS E SILVA, Rosa Virgínia. *O português arcaico: fonologia, morfologia e sintaxe*. São Paulo: Contexto, 2006.

WELKER, Herbert. *Dicionários uma pequena introdução à lexicografia*. Brasília: Thesaurus, 2004.