

Maria da Graça Krieger
mkrieger@unisinos.br

Tipologias de dicionários: registros de léxico, princípios e tecnologias

Dictionary typologies: lexicon records, principles and technologies

RESUMO - Este artigo contextualiza o perfil aplicado e teórico da lexicografia, compreendida como área da Linguística Aplicada. Destaca a complexidade da produção de dicionários e a existência da grande tipologia de obras de caráter dicionarístico. Refere também as diferenças metodológicas de suas composições e a influência das tecnologias informáticas na produção lexicográfica. Analisa as características gerais e finalidades do chamado dicionário padrão de língua, a mais prototípica das obras lexicográficas e a de dicionários com enfoques especiais na morfologia e na etimologia.

Palavras-chave: lexicografia, tipologia lexicográfica, tecnologia, *corpus* lexicográfico, dicionário geral de língua, dicionário etimológico, dicionário morfológico.

ABSTRACT - This article contextualizes the applied and theoretical profile of Lexicography as an area of Applied Linguistics. It highlights the complexity of the dictionary production and the existence of a great dictionary typology. It also studies the methodological differences of its composition and also the influence of information technology on the dictionary production. It analyses the general characteristics and the purpose of the monolingual dictionary, the most prototypical kind of work produced by Lexicography and also of dictionaries with special focus on morphology and etymology.

Key words: lexicography, dictionary typology, technology, lexicographic corpus, monolingual dictionary, etymological dictionary, morphological dictionary.

Cada projeto de dicionário é único e fala por si mesmo sobre seu conjunto de regras específicas, mas o vasto alcance das tarefas demanda uma organização rigorosa de fazer o melhor uso possível das fontes e das equipes. A crença comum de que a elaboração de um dicionário começa com a definição de palavras é tão ingênua como a idéia de que a construção de um prédio começa com a compra de materiais de construção (Landau, 2001, p. 343).

Introdução

A lexicografia é uma área de saber, cuja identidade está relacionada à produção de dicionários. Esta sua face aplicada é milenar, posto que o mundo antigo inaugurou a prática de relacionar palavras e sentidos para atender a necessidades de informação das coletividades lingüísticas. É assim, por exemplo, que nascem glossários na Grécia Antiga, nos quais eram listadas e definidas palavras de difícil compreensão de obras literárias. Estas, quando organizadas alfabeticamente ao final dos textos, constituíam os glossários, que representam as primeiras formas de manifestação lexicográfica.

Conta a tradição que a *Ilíada* e a *Odisséia* de Homero necessitaram de glossários para facilitar sua leitura. Assim também na Idade Média, quando o latim, que passou a ser

chamado de vulgar, já apresentava muitas diferenças com o latim clássico, que permanecia como a língua da liturgia, do direito e da universidade, foi necessário explicar o significado das palavras difíceis por meio de glosas, ou seja, de notas explicativas. É, portanto, desde tempos remotos que as civilizações de cultura cultuam a prática lexicográfica, dando vazão a um vasto universo tipológico de obras, denominadas, nem sempre apropriadamente, de glossários e dicionários.

No entanto, aquelas que, como um modelo canônico, sistematizam o léxico de um idioma, trazendo informações gramaticais e semânticas sobre as palavras de uma língua, integram o domínio da lexicografia propriamente dita, ou lingüística, que se constitui no objeto central deste artigo.

Abordar aqui o tema da lexicografia, mesmo que limitado a alguns aspectos relacionados a tipologias de

dicionários expressa uma dupla motivação. A primeira delas está relacionada à compreensão de que o dicionário é um instrumento de importância vital para as sociedades de cultura, já que é o único lugar que contém o léxico de um idioma; mas, contraditoriamente, é ainda um tipo de obra pouco estudada, mostrando que há ainda grande carência de estudos lexicográficos em nosso meio.

O dicionário é uma obra que nasce para atender a necessidades específicas das coletividades lingüísticas. Em primeiro plano, permite que elas tenham à disposição o registro do léxico de sua língua, numa correspondência com os significados que os recobrem. Em consequência, constitui-se em fonte de consulta sobre palavras, expressões, termos e sentidos desconhecidos:

[...] especificamente o ato de pergunta e de resposta acerca do significado de um signo revela seu caráter social, como um fenômeno distinguido pela sociedade entre as múltiplas ações que se orientam para o entendimento intersubjetivo; como um verdadeiro gênero da significação (Lara, 1996, p. 102).

Esta visão de Lara corresponde ao chamado dicionário de língua, sempre o parâmetro primeiro de todo o pensamento sobre o fato dicionário. Esse tipo costuma ultrapassar as informações semânticas, oferecendo informações gramaticais e lingüísticas. Como tal, descreve diferentes realizações das unidades lexicais, através do registro dos usos lingüísticos diferenciados que caracterizam as variedades regionais, as diacrônicas, bem como aquelas relacionadas aos usos e significados próprios das áreas científicas e técnicas. Assim, direta ou indiretamente, a obra lexicográfica traz informações funcionais e, por vezes, históricas sobre vários componentes dos sistemas lingüísticos. Outras vezes, vale-se de frases cotidianas e passagens literárias para exemplificar e abonar determinados usos. Esses são os dados mais comuns numa obra que por, traçar descritivamente, um panorama geral das realizações e virtualidades dos itens léxicos de um idioma, assumiu o caráter de referência sobre o léxico e seus modos de funcionamento em discurso.

Tal caráter está associado ao fato que o dicionário é o lugar formal e unitário de registro do componente léxico de um idioma. Nessa medida, constitui-se em paradigma lingüístico modelar dos usos e sentidos das palavras e expressões de uma coletividade lingüística, desempenhando o papel de código normativo da língua. É nessa mesma esteira que o dicionário adquire o estatuto de instância de legitimação do léxico, passando então a funcionar como uma espécie de cartório de registros, é ele que concede à palavra sua certidão de nascimento e, dessa forma, institucionaliza o conjunto léxico das línguas. Por tudo isto, o dicionário goza de uma autoridade que não é menor

nas sociedades de cultura que, inclusive, o entendem como instrumento da “verdade lingüística”, logo, inquestionável.

Todos esses papéis são cumpridos sob a aparente simplicidade de uma lista alfabética, da qual cada palavra é o lema, ou a cabeça, do verbete. Talvez em virtude dessa configuração estereotipada, relacionada ao tradicional código lexicográfico, a sociedade costuma compreender o dicionário como resultado apenas de um saber-fazer pragmático que pressupõe o domínio de uma fácil técnica de compilação de palavras, cujos usos e significados já são consagrados. Essa equivocada compreensão ganha maior concretude na medida em que o dicionário aparenta ser um objeto somente representativo de um saber lingüístico coletivo, sem interferência do lexicógrafo.

Entretanto, a obra dicionarística não se resume a uma listagem, mas, como um texto, possui regras próprias de organização. Nessa regularidade organizacional, há projetos específicos definidos em conformidade com os fins visados pelo lexicógrafo que imprime suas marcas subjetivas e ideológicas em sua obra. Antes ainda, a elaboração de uma obra dessa natureza não se resume a uma tarefa mecânica de compilação, mas exige uma competência especial sobre os fatos lingüísticos e a metodologia desse fazer.

Na realidade, o fazer lexicográfico consistente é sempre de grande complexidade. A valorização da competência na produção de um dicionário é a segunda razão motivadora deste artigo, que busca também evidenciar o papel de um dicionário singular pela proposta e pelo conteúdo e, desse modo, prestar o devido reconhecimento a seus autores. Trata-se do Dicionário Morfossemântico da Língua Portuguesa,¹ recentemente concluído por Sebald Bak, contando com a colaboração de Lauro Dick em acuidado trabalho de consultoria e revisão na tarefa de registrar as palavras do português com base na sua formação morfológica e etimológica.

Esta obra resulta do projeto de retomar e atualizar os cinco volumes do *Dicionário Morfológico da Língua Portuguesa* (Heckler *et al*, 1984, p. 88). Entretanto, o novo Dicionário assumiu uma identidade própria, alcançando um grau de excelência que o qualifica como instrumento imprescindível ao estudo da língua. Mais adiante, voltaremos a este dicionário específico que contribui para confirmar que toda prática lexicográfica qualificada requer um paradigma teórico-metodológico para repertoriar os dados léxicos e definir o modo de tratamento dos dados.

Lexicografia: da prática à disciplina lingüística

Por força de sua tradição prática, a lexicografia tem sido alinhada ao domínio da lingüística aplicada. Bem entendido, este enquadramento disciplinar fundamenta-se

¹ O Dicionário Morfossemântico da Língua Portuguesa foi concluído em 2006, na UNISINOS. Com mais de 120.000 entradas, conta também com uma versão eletrônica.

no pressuposto de que a dicionarização do léxico, como já mencionamos, resulta da aplicação de um paradigma teórico baseado na lingüística. Tal aplicação responde, ao menos, entre os lexicógrafos lingüistas pela consciência de que uma lexicografia assim orientada deve atingir um patamar de cientificidade.

É ilustrativo desse fato o pensamento que enfatiza a necessidade de um conhecimento teórico para identificar as unidades léxicas dignas de registro e operar o tratamento semântico adequado de modo a oferecer a correta informação ao usuário. Assim, a lexicografia:

exige ao mesmo tempo um saber teórico (definição de unidades lexicais, tipologia das definições, e, na maioria dos casos, uma opção a favor desta ou daquela teoria semântica) que depende de uma semântica lexical (ou de uma lexicologia semântica) (Greimas e Courtès, s.d., p. 256).

De fato, a ciência da linguagem tem oferecido fundamentos para o registro e equacionamento dos itens léxicos. Nessa medida, a lexicografia é particularmente tributária da lexicologia e da semântica para equacionar o plano lexical e o das acepções e definições das palavras. De igual modo, não pode dispensar estudos sociolingüísticos e discursivos para descrever o funcionamento do léxico, entre muitas outras áreas do conhecimento lingüístico. A crescente interface com a lingüística de *corpus* é também comprobatória da busca de bases operacionais e metodológicas da lexicografia contemporânea.

Por sua vez, estudos sistemáticos sobre distintos componentes lingüísticos e pragmáticos das produções lexicográficas, realizados à luz de fundamentos da ciência da linguagem, inscrevem-se no campo da lexicografia teórica. E, de modo especial, no da metalexicografia, domínio de saber que visa a estudar os problemas referentes à produção de dicionários lingüísticos.

Teoria lexicográfica, lexicografia teórica ou metalexicografia são as denominações mais usuais para esse componente teórico da lexicografia, que muitos autores começam a empregar para diferenciar essa dimensão teórica tanto da prática concreta ou confecção de dicionários, como de âmbito disciplinar também afim, mas claramente diferenciado por seus objetivos e métodos, da lexicologia (Fernández, 2003, p. 36).

Isso corresponde a estabelecer uma teoria lexicográfica com perfil autônomo no sentido de definir seu objeto, postular princípios e metodologias e identificar os problemas envolvidos nas aplicações lexicográficas. Estes conhecimentos devem contribuir para que um objeto cultural, como o dicionário, atinja um desejável padrão de qualidade.

Por outro lado, o desenvolvimento dos estudos lexicográficos teóricos tem ampliado as condições de análise crítica dos dicionários, permitindo aprofundar a compreensão de que eles não são todos iguais, nem tampouco são neutros como a sociedade costuma pensar. Ao contrário, há grandes diferenças tipológicas, em razão dos

objetivos de cada um, das necessidades dos usuários previstos, da amplitude do repertório léxico, da escolha das informações registradas, entre outros aspectos. As diferenças ocorrem, inclusive, dentro uma mesma tipologia, já que nem todos apresentam o mesmo, nem as mesmas definições a comprovar que os dicionários, ao modo dos textos e discursos, contêm as marcas de seu produtor, embora estas não sejam tão evidentes, ficando diluídas sob os cânones lexicográficos. As diferenças incidem, portanto, sobre os aspectos qualitativos, seja pelas propostas diferenciadas, seja pela acuidade do fazer, fator que diferencia as obras confiáveis daquelas que não merecem a confiança que a sociedade costuma conferir aos dicionários.

Sem adentrar nesse mérito, passamos a reter a atenção nas tipologias dicionarísticas, salientando alguns princípios e aspectos determinantes da organização dos dicionários lingüísticos. Tal como entendemos, integram a lexicografia lingüística as obras que repertoriam léxicos, independente de ângulos que cada proposta prioriza. Em contraponto, alguns estudiosos postulam a dicionarística como uma área que cobre outras proposições temáticas de sistematização de informação, organizadas ao modo de dicionários ou assim denominadas.

Da tipologia de dicionários

Diante da amplitude do tema, privilegiamos alguns aspectos do universo da prática e da metodologia referentes à produção de dicionários que, longe de ser uniforme, apresenta uma grande variedade tipológica – dicionário monolíngüe, bilíngüe, dicionário geral, tipo thesaurus, tipo padrão, de usos, minidicionário, dicionário escolar, – entre tantas outras possibilidades. Isto para ficar no âmbito das obras de referência lingüística, ou seja, as que registram o léxico de forma sistemática e são, conseqüentemente, consideradas como paradigmas lingüísticos, independente de sua extensão.

No universo dos dicionários lingüísticos, encontram-se também aqueles que tratam de temáticas específicas, embora com funções distintas: verbos e regimes, sinônimos, antônimos, morfológicos, etimológicos, além de outros tantos casos.

Junto à diversidade tipológica, vale observar que os critérios classificatórios são também variáveis, explicando denominações que, em geral, dependem do componente tomado por base a exemplo dos dicionários técnico-científicos ou terminológicos que repertoriam os termos técnico-científicos de alguma área de conhecimento. Essas produções que, muito comumente, também aparecem sob a forma de glossários, restringem-se a repertoriar o que é convencionalmente chamado de léxico especializado. Tal restrição, por sua vez, explica a oposição ao dicionário geral, denominação, muitas vezes aplicada em razão da obra cobrir a totalidade das realizações léxicas de um

idioma. Neste caso, a totalidade não significa o registro exaustivo das palavras de uma língua, desde as mais antigas até os neologismos mais recentes, até porque a lexicografia não consegue acompanhar o dinamismo lexical; mas caracteriza a abrangência do componente léxico, sem privilegiar uma temática específica. A especificidade está nos dicionários que têm por objeto um subconjunto léxico, por exemplo, a terminologia da química, da informática ou do meio ambiente.

Outra diferença reside no fato de que os dicionários gerais, monolíngües ou não, cobrem a totalidade da língua tendo por base o critério da frequência de uso da palavra ou da expressão; em contraponto, os especializados selecionam seus objetos com critérios específicos. Nesse sentido, os terminológicos registram os termos de uma área, considerando sua importância conceitual. Assim, embora considerem também a frequência de uso, o conceito é o princípio chave. Por esses direcionamentos, a lexicografia assume caráter semasiológico, porquanto se orienta do significante para o significado; em sentido inverso, a terminografia, ou seja a área de produção de obras terminológicas, fundamenta-se na onomasiologia, selecionando os itens lexicais especializados a partir dos conceitos que veiculam. É também em razão disso que tais obras não abrangem todas as classes gramaticais, já que as palavras gramaticais, como artigos, conjunções e preposições, não têm valor conceitual, mas função relacional na construção das frases.

O conjunto dos registros de entrada define a macroestrutura dos dicionários ou a sua nomenclatura. A extensão dessa nomenclatura tem sido tomada como um parâmetro tipológico, bem como responde por denominações como dicionário do tipo “thesaurus”, tipo padrão ou minidicionários. Em relação a estes últimos, a nomenclatura é reduzida, embora não haja clareza sobre os critérios de seleção lexical. Costumam ter entre 20.000 e 35.000 entradas no caso do Brasil, em contraponto aos de maior cobertura, que alcançam cerca de 200.000 entradas. A falta de informações explícitas nas obras brasileiras seja sobre os números reais, seja sobre os critérios de constituição da nomenclatura dificulta uma categorização mais específica das obras que circulam em nosso meio.

Entretanto, menos do que a extensão exata, importa-nos mais as características tipológicas. Nesse sentido, a denominação “thesaurus” está vinculada à idéia de exaustividade histórica de registro, cobrindo desde palavras antigas aos modernos neologismos. Este modelo chega alcançar algo em torno de 400.000 verbetes em países de grande tradição lexicográfica como a França, a Espanha e a Inglaterra.

É interessante também observar que este modelo de larga cobertura lexical reflete uma concepção muito clássica de que o dicionário é o tesouro da língua. Ou melhor, é o lugar que guarda o tesouro da língua, cumprindo um papel de memória diante do dinamismo do léxico que se transfor-

ma ao longo do tempo, seja pelo acolhimento de novas palavras e sentidos que surgem, seja pelas muitas que caem em desuso. De certo modo, a concepção de guardar o tesouro da língua está também relacionada à etimologia da palavra dicionário, o sufixo *arium* significando depósito, indica lugar onde as palavras são guardadas.

Os dicionários do tipo clássico possuem ainda outras características como a presença em suas nomenclaturas de palavras usadas apenas por escritores, costumando recorrer a citações literárias, com valor de abonações. Assim, junto à exaustividade dos registros léxicos, procuram selecionar as palavras que representam o “melhor uso” da língua, justificando a escolha do texto como modelo da mais acabada expressão lingüística. Daí também a razão pela qual o dicionário é compreendido como o lugar que registra o “melhor da língua”, estabelecendo o padrão modelar de expressão das comunidades lingüísticas e reafirmando seu papel de código normativo das regras do bem-dizer.

Dessa forma, a constituição tradicional da nomenclatura na obra lexicográfica monolíngüe, em geral chamada de dicionário de língua, é de caráter qualitativo e não apenas quantitativo. Tal modelo é geralmente tomado como paradigma, por excelência, do fato dicionário, já que sua proposta é a de repertoriar o conjunto das palavras e expressões de uma língua e oferecer inúmeras informações sobre a gramática da palavra-entrada, seus sentidos e seus usos. O modelo completo cobre ainda fonética, a história, via etimologia, os dados do funcionamento lingüístico e discursivo das palavras-entrada. Por tudo isto, é uma espécie de dicionário padrão. O conjunto de informações que este tipo de dicionário oferece o tornam um lugar privilegiado de lições sobre a língua (Krieger, 2003, p. 70-87).

Não é intenção agora avançar nas características dessa tipologia, mas salientar que, para cumprir o programa de repertoriar as unidades lexicais, definir-lhes o significado e configurá-las descritivamente, a prática lexicográfica necessita fazer frente à heterogeneidade constitutiva do léxico. Este não é um bloco monolítico, apesar de que o dicionário projeta a imagem do léxico como um componente linear e formalmente pré-estabelecido. Essa imagem contribui para reforçar a idéia comum de que a lexicografia limita-se a uma atividade pragmática compilatória, cabendo ao lexicógrafo apenas reproduzir o dito e agregar informações gramaticais e semânticas, entre outras que considerar relevantes. Em outras palavras, é como se as entradas, registradas em cada obra e as definições de cada item estivessem “prontas”, à espera de alguém que goste de “palavras” e as coloque num dicionário.

Mas, efetivamente, o léxico de uma língua é um conjunto heterogêneo sob alguns ângulos de sua composição e o lexicógrafo deve ser capaz de identificar sua composição. Dentre os fatores que determinam a variada formação do conjunto léxico de um idioma, destacam-se: o tempo, o espaço e o registro. É em razão da variação diacrônica que o léxico geral de um idioma contém pala-

bras antigas, que caem em desuso. No entanto, coexistem com os neologismos sob o prisma do conjunto do acervo léxico de uma língua. A variação diatópica relaciona-se à integração no mesmo conjunto lexical de algumas palavras distintas, mas usadas com sentido igual, caracterizando o falar de diferentes regiões geográficas. A variação denominativa ocorre tanto dentro de um mesmo país, quanto na relação com outra nação que pratica o mesmo idioma a exemplo do que ocorre entre Portugal e Brasil.

Já a variação diastrática está relacionada à compreensão de que a constituição do léxico é também resultante de estratos ou faixas horizontais, as quais respondem pelos usos diferenciados das palavras, determinando registros como: culto, popular, literário, vulgar, coloquial. Incluem-se aí o registro dos sentidos de domínios do saber especializado como a botânica, a física, as tecnologias digitais, entre outros.

Estes componentes ilustram a heterogeneidade constitutiva do léxico, que, em primeiro plano, cumpre a função de nomear os seres, os objetos, as ações e processos que identificam o mundo fenomenológico e aquele percebido pelos homens. Ao mesmo tempo, tal função explica o fato do componente lexical transformar-se continuamente. Como já disse Horácio na *Arte Poética*:

Muitas palavras que já morreram terão um segundo nascimento, e cairão muitas das que agora gozam das honras, se assim o quiser o uso, em cujas mãos está o arbítrio, o direito e a lei da fala (1984, versos 70-72).

Toda essa ordem de fatores revela o léxico é um componente multifacetado em constante mobilidade; conseqüentemente, funcionando como o pulmão das línguas. Dessas características resulta o entendimento corriqueiro, mas que também influenciou a lingüística, de que este componente é inorgânico, descontínuo e idiossincrático, embora seja altamente sistêmico como exemplarmente atesta a morfologia derivacional.

Lexicografia e tecnologias informáticas

As tecnologias informáticas exercem também influência na estruturação dos dicionários. Em decorrência, há uma fase nova na história do fazer dicionarístico, considerando dois fatores essenciais: o princípio de seleção lexical e o aporte decorrente das tecnologias informáticas. Os diferenciais são então de natureza epistemológica e pragmática, incidindo, igualmente, sobre a escolha das fontes que servem de base à seleção lexical.

Se em seus moldes clássicos, a lexicografia buscou registrar o componente léxico ao longo de sua evolução, considerando fases antigas da história da língua e ainda fazendo escolhas para representar o bom uso das palavras, ela altera-se na medida em que passou a tomar por objeto as vozes de uso contemporâneo da língua, sem julgamentos pré-estabelecidos. Resulta daí:

uma lexicografia baseada fundamentalmente na comunicação, que parte do valor intrínseco do vocabulário no processo comunicacional, dos modos de uso e das situações de uso de uma unidade léxica dentro de uma coletividade lingüística (Haensch, 1982, p. 19).

Dessa forma, são definidos os novos fundamentos epistemológicos que levam a priorizar cortes sincrônicos, dando lugar ao surgimento dos chamados dicionários de uso, nos quais a nomenclatura é estruturada com os itens lexicais realizados e falados com frequência. Os percentuais de frequência seguem parâmetros gerais determinados pela e para a lexicografia dessa natureza que se caracteriza, igualmente, pela grande preocupação de constituir um *corpus* representativo dos enunciados lingüísticos, constituído pela maior diversidade possível de gêneros textuais. Para tanto, são incluídas manifestações orais de forma a oferecer um conjunto de entradas que reflita o léxico real e não virtual, como assinala Borba (2003, p. 18). Explica-se, nesse contexto, a relevância atribuída ao grande número de ocorrências das unidades lexicais nas situações comunicativas, condição para validar o *corpus* selecionado e os resultados da pesquisa.

No estudo da linguagem, um *corpus* é qualquer conjunto de texto coletado com o objetivo de analisar suas características. Atualmente quando se fala de um *corpus* em lexicografia, é entendida a referência a um *corpus* eletrônico, geralmente contendo um vasto número de palavras das mais diferentes fontes (Landau, 2001).

Por outro lado, a constituição dos *corpora* informatizados agilizou o trabalho lexicográfico, que tem se beneficiado largamente do grande número de aplicativos que identificam rapidamente as palavras e seus percentuais de ocorrência. Em síntese, os princípios de seleção lexical, a escolha das fontes de coleta e o apoio logístico das tecnologias informatizadas identificam a nova fase da lexicografia que, na busca de cientificidade, mantém um diálogo incessante com a Lingüística de *Corpus*.

No bojo dessas transformações, e do advento dos dicionários de uso, é interessante observar que a norma lingüística, ou o dizer “correto” não está mais vinculada a um padrão modelar de expressão, mas à aceitabilidade do funcionamento de um idioma, pois resultado da utilização recorrente de unidades lexicais em suas formas e sentidos por toda a comunidade lingüística. Dessa forma, neste tipo de dicionário, não se trata mais de registrar o “melhor” da língua, mas a língua “real”, tornando-se a efetividade do uso o novo parâmetro de referência lingüística. Mais uma vez, como disse Horácio, confirma-se o uso como o grande mestre.

Além desses aspectos que incidem sobre a estruturação de uma nova tipologia lexicográfica, os suportes informatizados acrescentaram categorias a esse universo com base nos suportes que os veiculam. É assim que, para Welker, “o termo dicionários eletrônicos refere-

se a dicionários: (1) usados no processamento computacional da linguagem natural; (2) em CD-ROM; (3) *online* (acessíveis na internet); (4) portáteis” (2004, p. 225).

As inovações tecnológicas de suporte têm favorecido a publicação de muitas obras, e facilitado os modos de busca. Alguns novos princípios organizacionais são introduzidos para uma consulta mais ágil, reafirmando a influência da tecnologia sobre o universo da lexicografia.

Etimologia e morfologia: a origem e a formação do léxico nos dicionários

Os elementos antes abordados evidenciam que os dicionários têm estruturas e metodologias distintas, e mais que isto, finalidades distintas. Entretanto, no quadro dos dicionários lingüísticos, o objeto comum a todos é sempre o componente léxico. Logo, variam apenas as aproximações ao léxico, em razão do perfil e dos fins de cada obra; em conseqüência, variam os enfoques em razão dos ângulos que cada obra destaca. De diferentes modos, é sempre uma contribuição para o conhecimento das palavras e expressões de uma língua.

Diante disso, vale observar a organização dos verbetes dos dicionários de língua, do tipo padrão. Estas microestruturas comportam muitas informações gramaticais e lingüísticas sobre a palavra-entrada, as mais completas trazem também dados etimológicos, aspecto que passa a reter nossa atenção.

A informação sobre o étimo significa recuperar a origem de uma palavra, compreendendo a base lexical antiga, ou seja, o componente morfológico a partir do qual gerou-se a palavra que se torna o objeto da consulta ao dicionário. Em Houaiss e Villar, (2001), cuja obra caracteriza-se pelo registro sistemático do étimo, lê-se que: “Étimo, do grego *Étumon*, comporta uma outra acepção, qual seja: ‘o verdadeiro significado da palavra segundo sua origem’.

Dois sentidos aliam-se, portanto, na definição de étimo: (1) informação sobre as bases lexicais que deram origem à palavra que passou a integrar o vocabulário de um idioma; (2) a verdade sobre esta palavra.

Em relação a este último sentido, é interessante observar o interesse constante do homem sobre a história da palavra, como componente da “verdade” do seu sentido. Do ponto de vista etimológico, esta verdade está associada à busca da razão das denominações escolhidas pelo homem ao cunhar os itens lexicais de seu idioma.

O registro do étimo não se restringe aos dicionários de língua de modelo canônico, mas também está presente em dicionários específicos em que se incluem os morfológicos e os etimológicos propriamente ditos. Nos dicionários gerais, a etimologia está a serviço da lexicografia; ampliando as informações sobre a palavra, descrevendo sua formação histórica.

Entretanto, nos específicos, caso dos etimológicos, essas informações constituem sua espinha dorsal. Vale

dizer, neles, as nomenclaturas fundam-se sobre as palavras que possuem em sua constituição morfológica um componente de base transladado de outra língua. Já as microestruturas são variáveis, dado que podem conter a data do primeiro registro do vocábulo na língua em foco, além da relação das palavras cognatas, entre outras informações sobre a história e a composição dos vocábulos. Tais dados caracterizam ainda a natureza diacrônica de uma obra de étimos, porquanto seu objetivo maior é registrar a evolução do vocabulário através do tempo, resgatando a “suposta pré-história das palavras”, como diz Haensch (1982).

O estudo da trajetória dos itens léxicos reveste-se de uma complexidade incomum e requer uma competência cognitiva ímpar:

O etimólogo, em efeito, por mais modesto que seja seu campo de ação, isto é, ainda que somente se proponha a explorar uma só língua, necessita manejar com soltura o vocabulário das línguas irmãs com sua respectiva dialetologia, tendo que conhecer a gramática histórica de cada uma, suas peculiaridades fonéticas nas várias etapas evolutivas, as relações de toda a família com o tronco comum e necessita localizar no tempo e no espaço os contatos e intercâmbios das atividades e culturas dos povos que se servem das línguas relacionadas (Casares, 1992, p. 32).

Esse conjunto de competências está refletido no Dicionário Morfossemântico da Língua Portuguesa, referido no início. Embora predomine sua feição morfológica, os verbetes são organizados, considerando a base etimológica da palavra-entrada, como se pode observar no Quadro 1.

Quadro 1. Estrutura do Dicionário Morfossemântico da Língua Portuguesa.

Dosagem

'Relativo a quantidade'.

dosagem	dos-ag-em	S
dosar	dos-a-r	V
dose	dos-e	S
doseamento	dos-e-a-ment-o	S
dosear	dos-e-a-r	V
dosificar	dos-i-fic-a-r	V do latim <i>facere</i> 'fazer'.
dosimetria	dos-i-metr-i-a	S do grego <i>metron</i> 'medida'.
dosimétrico	dos-i-métr-ic-o	A

Origem: do grego *dósis* 'quantidade de medicamento ministrada ao doente', derivado *dedídomi* 'dar'. Veja a família *dação*.

A organização dessa microestrutura demonstra a íntima relação com os objetivos que nortearam a elaboração da obra:

- facilitar o estudo da morfologia da Língua Portuguesa;
- facilitar o estudo de aspectos morfossemânticos pela disposição das palavras em famílias lexicais, dando conta, portanto, dos vocábulos cognatos.

O verbete que se apresenta no Quadro 2 confirma o intuito e a configuração didática da obra.

Quadro 2. Disposição em famílias lexicais no Dicionário Morfossemântico da Língua Portuguesa.

Mania

'Doença mental caracterizada por um estado de excitação do espírito; esquisitice; excentricidade'.

esmaniado	es-man-i-ad-o	A
esmaniar	es-man-i-a-r	V
mania	man-i-a	S
maniaco	man-i-ac-o	A
manicômio	man-i-côm-io	S do grego <i>komêin</i> 'cuidar'.
manigrafia	man-io-graf-i-a	S do grego <i>gráphein</i> 'escrever'.

Origem: do grego *mania* 'loucura, demência, humor, sombrio, delírio profético, loucura de amor', do verbo *mainein* 'tornar louco', com o lexema indo-europeu **men-*, confira *mémona*, de **ménein* 'desejar'; *ménos* 'mente'; no latim *mente*.

Tal como se observa, a morfologia e a etimologia das unidades lexicais se implicam mutuamente, aspecto que as famílias de palavras reiteram de modo incontestável, já que possuem a mesma base formadora. A articulação entre as duas áreas confere a este dicionário a eficiência didática aliada à cientificidade dos registros etimológicos nos moldes da tradição de pesquisa na área.

Vale ainda remarcar que dicionários dessa natureza não se limitam a entradas com cognatos, mas palavras isoladas também merecem o devido registro, salientada sua composição morfológica, conforme ilustram *diapasão* e *jargão* (Quadro 3).

Quadro 3. Exemplo de palavras isoladas no Dicionário Morfossemântico da língua Portuguesa.

Diapasão di-a-pas-ão S 'Instrumento usado para regular a altura dos sons'.

Origem: do grego *diá pasón chordón* 'através de todas as cordas da oitava musical', através do latim *diapaason* 'escala musical de sete tons'.

Jargão jarg-ão S 'Linguagem viciada, imperfeita'.

Origem: do francês antigo *jargon* 'língua estrangeira e ininteligível'.

A organização de todos os verbetes, mesmo os das "palavras isoladas" revelam a preocupação em evidenciar a forma de constituição de cada item léxico, vale dizer, com os mecanismos de produtividade lexical. Novas palavras são formadas por meio de determinados padrões, como o uso de prefixos e sufixos agregados a um radical entre outros processos que permitem criar e entender a

composição formal das palavras. Com esse quadro de informações, o consulente pode, com facilidade, identificar os processos de formação das palavras de uma língua. Em especial, os estudos de morfologia derivacional podem muito se beneficiar desses conhecimentos.

O conteúdo dos vocábulos e as correlações semânticas entre os cognatos também estão projetados nesse tipo de obra que, entre tantos méritos, permite evidenciar que o componente léxico de um idioma possui regras orgânicas de formação. Em consequência, auxilia a reverter a equivocada idéia de que o léxico é inorgânico e idiossincrático. Revela-se aqui de mais uma, dentre as muitas funções sócio-cognitivas que os dicionários morfológicos e etimológicos oferecem a seus consulentes.

Além disso, toda essa organização confirma que os fins de cada obra e a consideração pelo usuário visado são fatores que influem largamente na configuração dos moldes organizacionais dos dicionários. Essa relação agrega-se ao quadro de fatores que determinam e explicam o universo de tipologias de dicionários, bem como revelam a complexidade que caracteriza a engenharia lexicográfica.

Referências

- BORBA, F.S. 2003. *Organização de dicionários: uma introdução à lexicografia*. São Paulo, Editora UNESP, 356 p.
- CASARES, J. 1992. *Introducción a la lexicografía moderna*. 3ª ed., Madri, Consejo de Investigaciones Científicas, 354 p.
- FÉRNANDEZ, D.A. 2003. La lexicografía como disciplina lingüística. In: A.M.M. GUERRA (org.), *Lexicografía española*. Barcelona, Editorial Ariel, p. 31-52.
- GREIMAS, A.J. e COURTÈS, J. s.d. *Dicionário de Semiótica*. São Paulo, Cultrix, 492 p.
- HAENSCH, G.; WOLF, L.; ETTINGER, S. e WERNER, R. 1982. *La lexicografía: de la lingüística teórica a la lexicografía práctica*. Madrid, Gredos, 590 p.
- HECKLER, E., BACK, S., MASSING, E.R. 1984. *Dicionário morfológico da língua portuguesa*. 5 v., São Leopoldo, Unisinos.
- HORÁCIO. 1984. *Arte poética*. Lisboa, Inquérito.
- HOUAISS, A. e VILLAR, M. 2001. *Dicionário Houaiss da Língua Portuguesa*. Rio de Janeiro, Objetiva, 2922 p.
- KRIEGER, M.G. 2003. Dicionário de língua: um instrumento didático pouco explorado. In: C.S. TOLDO (org.), *Questões de Lingüística*. Passo Fundo, Universidade de Passo Fundo, p. 70-87.
- LANDAU, S.I. 2001. *Dictionaries: the art and craft of lexicography*. 2ª ed., Cambridge, Cambridge University Press, 477 p.
- LARA, L.F. 1996. *Teoría del diccionario monolingüe*. Mexico, El Colegio de México; Centro de Estudios Lingüísticos y Literarios, 273 p.
- WELKER, H.A. 2004. *Dicionários: uma pequena introdução à lexicografia*. Brasília, Thesaurus, 287 p.

Submetido em: 10/2006

Aceito em: 11/2006

Maria da Graça Krieger

Professora doutora do Programa de Pós-Graduação em Lingüística Aplicada
UNISINOS, RS, Brasil