

ANÁLISE DE DOIS DICIONÁRIOS GERAIS DO PORTUGUÊS BRASILEIRO CONTEMPORÂNEO: O AURÉLIO E O HOUAISS

Maria Tereza Camargo Biderman*

RESUMO: Este artigo faz uma avaliação de dois dicionários gerais do português recentemente publicados: o Dicionário HOUAISS da Língua Portuguesa (2001) e a última versão do AURÉLIO Século XXI (2000). Examina-se primeiro a fundamentação teórica dessas obras, evidenciando-se as falhas existentes com respeito a conceitos e métodos da Teoria Lexical, mostrando-se como é imprescindível que dicionários se baseiem em um *corpus* de textos já realmente produzidos. Examina-se também a abrangência e a representatividade do léxico registrado nestas obras e os critérios usados na seleção das palavras e dos lemas da nomenclatura, assim como o tratamento dado aos verbetes nestes dicionários.

Palavras-chave: dicionário, dicionário geral, lexicografia, *corpus* informatizado, macroestrutura e microestrutura do dicionário.

1. INTRODUÇÃO

O léxico de uma língua natural constitui uma forma de registrar e armazenar o conhecimento do universo. Como diz Alan Rey, no prefácio do Dicionário da língua francesa – *Petit Robert*, ed. de 1994 – o léxico vem a ser a medida de tudo, pois exprime e reflete o universo cultural da sociedade. Ora, o tesouro vocabular de um idioma constitui um conjunto cujas dimensões não é possível precisar. De fato, o número total de palavras de uma língua de civilização é indeterminado;

* UNESP, Faculdade de Ciências e Letras, Campus de Araraquara.

talvez atinja uns 500.000 vocábulos ou mais. Além disso, sendo um conjunto aberto, novas palavras são criadas continuamente pelos usuários, especialmente os mais cultos e mais criativos, e também pelos cientistas (as terminologias científicas).

Por outro lado, como afirma Lara (1992, p. 20), “o dicionário representa a memória coletiva da sociedade e é uma de suas mais importantes instituições simbólicas”. Assim sendo, o dicionário vem a ser o depositário do acervo lexical da cultura. E como diz Alan Rey, o dicionário é a memória lexical de uma sociedade; constitui o acervo e o registro das significações que nossa memória não é capaz de reter (Prefácio do *Petit Robert*).

O dicionário é também e sobretudo um produto lingüístico; constitui “o resultado de uma infinidade de atos verbais que, na experiência social, desligaram-se de seus atores” para passar a fazer parte do patrimônio cultural coletivo, especialmente o que foi dito inteligentemente no seio dessa sociedade. (Lara, 1992, p. 20). O conjunto dos usos sociais da língua é refletido pelo dicionário.

Além disso, o dicionário descreve o léxico em função de um modelo ideal de língua – a língua culta e escrita. Só circunstancialmente registra os padrões subcultos, ou desviantes da norma padrão, tais como os usos dialetais, populares, giriáticos.

Dessa forma o dicionário convalida e promove a linguagem aceita e valorizada em sua comunidade.

2. FUNDAMENTAÇÃO TEÓRICA

2.1 Conceito de unidade léxica e identificação das unidades complexas.

Um dicionário precisa ser fundamentado em uma *teoria lexical*, levando em consideração premissas básicas da lexicologia. Como, por exemplo, o conceito de unidade léxica. Muitas vezes é extremamente difícil aplicar esse conceito, já que, na prática discursiva, é complexo delimitar as unidades lexicais no contexto. E não só é

difícil identificar a unidade léxica como é complicado eleger o lema para ser o *caput* do verbete.

De fato, a identificação da unidade lexical ao nível do discurso constitui um grande desafio para um lexicógrafo, pois as fronteiras entre as palavras muitas vezes são difusas. É fato inconteste que as fronteiras entre uma *unidade lexical complexa* e um *sintagma discursivo* são fluidas. Existe toda uma gama de graus de lexicalização entre os elementos de uma combinatória lexical – dos sintagmas cristalizados como lexias compostas e/ou complexas às unidades fraseológicas, às *expressões idiomáticas*. Essas unidades complexas, contudo, têm uma coesão interna do ponto de vista semântico, fato esse que deveria levar o dicionarista a individualizá-las como unidades já consolidadas do léxico. Nesse caso essa unidade complexa deveria constar como entrada do dicionário; a hipótese contrária implicaria a inclusão de tais combinatórias ou fraseologias como subentradas no interior de um verbete, cabendo também ao dicionarista decidir onde elas melhor se encaixariam.

Conseqüentemente, o conceito de unidade léxica do dicionarista reflete-se na organização da *macroestrutura do dicionário*, bem como os critérios por ele usados na seleção dos lemas. Ora, o lema representa um lexema. E o lexema é uma entidade abstrata que se manifesta no nível do discurso de modo bem diversificado por ser o português uma língua flexiva. Por outro lado, nas realizações discursivas, as fronteiras entre uma unidade lexical complexa e um sintagma discursivo livre são muito difusas, exigindo do lexicógrafo uma boa formação teórica para poder decidir entre casos limítrofes. Mais ainda: o conceito de unidade lexical levanta problemas teóricos com conseqüências práticas na sua identificação e tratamento ortográfico e lexicográfico.

2.2 Vejamos como operaram estes dicionários

De um modo geral, tanto o AURÉLIO como o HOUAISS revelam desconhecimento da Teoria Lexical, Gramatical e Lingüística.

Convém lembrar que a versão do AURÉLIO Século XXI coincide, em grande parte, com a edição de 1986, publicada ainda em vida de Aurélio Buarque de Hollanda Ferreira. Estruturalmente não muda quase nada; há acréscimos de verbetes mas não em volume substancial.

Na seção “como usar o AURÉLIO Século XXI” são expostos parâmetros de instrução ao consulente que refletem os conceitos e métodos subjacentes à elaboração do dicionário. No caso do HOUAISS tais especificações e instruções também precedem o dicionário numa seção denominada “Chave do dicionário”, texto assinado por Mauro de Sales Villar.

Nesta seção do AURÉLIO são classificadas como locuções (cf. nota 19) segmentos que são, de fato, *unidades léxicas complexas*, típicas das línguas de especialidade científica. O exemplo dado *resistência dos materiais* não é uma locução mas um termo da Físico-química e da Engenharia. Aliás, é típico dos domínios científicos e técnicos os conceitos serem expressos por unidades fraseológicas que constituem uma unidade léxica do ponto de vista conceitual, sendo esse exemplo um caso de composição. O processo de composição é recurso muito freqüente na geração de termos das ciências e das técnicas. Inversamente, constituem locuções propriamente ditas: 1) locuções adverbiais – *a baila, a sério, em pé*; 2) locuções prepositivas – *além de, até a, diante de, por volta de*; 3) locuções conjuncionais – *ainda que, como se, de modo que, por conseguinte*, etc. O AURÉLIO inclui estas unidades complexas como subentradas da palavra-base sem classificá-las, não revelando, pois, a consciência de que se está diante de unidades complexas tradicionalmente classificadas como locuções gramaticais.

Já o HOUAISS explica: “A unidade lexical pode ser uma palavra simples, uma palavra composta por hífen, uma locução, uma redução (do tipo símbolo, sigla, abreviatura, etc.), pode ainda ser múltipla e, em casos mais raros, tratar-se de pequena frase”. No item 42 e seguintes deste prefácio Villar generaliza o uso do termo técnico *locução*, atribuindo-lhe o sentido de “unidade complexa do léxico”,

o que não se pode aceitar. Infelizmente esta introdução muito minuciosa, em que Villar buscou explicitar todas as categorias e práticas do dicionário HOUAISS, mostra que seu autor não está familiarizado com as muitas obras que tratam das ciências do léxico, particularmente da morfologia lexical e dos processos de formação de palavras, para não falarmos de outras complexidades como a história do léxico. Também se deveria ratificar o termo *nominata* usado por Villar, pois o termo consagrado em Lexicologia/Lexicografia é *nomenclatura*. Os teóricos do léxico também usam o termo *macroestrutura* que refere, porém, um conceito ligeiramente distinto.

2.3 Consideremos unidades menores que a palavra: *morfemas derivacionais e elementos de composição*. Até alguns anos atrás os dicionários não costumavam incluir esses elementos mórficos. De fato, não constituem unidades integrantes do patrimônio lexical. É, porém, útil para o consulente a inclusão destes formantes do léxico; eles podem servir não só à criação de palavras novas como também a um melhor entendimento da constituição do léxico.

Na seção “como usar o AURÉLIO Século XXI” há alguns informes relativos aos formantes lexicais que são teoricamente incorretos. Cf. item “3. ? Indica elemento de composição: prefixo, sufixo, infixo (sic)”. Os prefixos e sufixos não são considerados na Teoria Lexical como elementos de composição, mas como afixos ou morfemas de derivação ou derivacionais. Um dicionarista precisa conhecer bem os processos de formação de palavras; aliás, a composição é um processo bem diverso da derivação.

Constata-se que estes dicionaristas carecem de competência especializada no que tange à questão da *formação de palavras* na língua. Por exemplo: inúmeras vezes as diversas classes de formantes que integram a morfologia lexical do português não são corretamente identificadas e tratadas. Assim, é discutível a inclusão inadequada de muitas unidades como *elementos de composição* em desacordo com a Teoria Lexical. Essa questão é particularmente séria no

HOUAISS, em que os deslizos são inúmeros, eivando todo o dicionário com disparates. De fato, muitas vezes são aí incluídos como *elementos de composição* pseudomorfemas que não podem ser assim categorizados.

No AURÉLIO, às vezes a classificação é correta e o tratamento do formante no verbete está correta, e, em outras, não. Isso porque é preciso distinguir um formante que, de fato, gera palavras novas no português de vocábulos emprestados pelo português ao latim e freqüentemente a alguma língua moderna (inglês ou francês) que, por sua vez, havia emprestado a palavra do latim. Veja-se, por exemplo, os verbetes: “*alti-* (do latim) = alto, elevado: *altiloquente, altimurado, altícola*”. Onde recolheram tais palavras? Não encontrei registro dessas unidades em nenhum texto do *corpus* da UNESP.¹ “*Digit(i)* [Do lat. *digitus*,i] El. comp. = ‘dedo’, ‘dígitto’: digitar, digitigrado...” Só *digitar* seria derivado dessa base latina. Mas, de fato, é empréstimo em português, e não, derivação. Já *prestidigitação*, citado a seguir, é empréstimo do francês e não derivação no português.

É o caso também do verbete *arci-* que registra *arcífero, arciforme* como derivados desse *soi-disant* elemento de composição. Mas esses cultismos não são palavras geradas no português, mas empréstimos. Aliás, esse é um problema recorrente – com freqüência se faz menção a produtos vocabulares que seriam criações vocabulares do português e que são empréstimos, ou tomados diretamente do latim, ou de uma língua moderna (inglês ou francês), idioma esse que tomou emprestada a palavra a uma língua clássica. Já o tratamento de *arqui-* está correto.

Se confrontamos a edição de 1986 do AURÉLIO com a versão XXI, vemos que o XXI altera o tratamento da matéria. Veja-se o

¹ O *corpus* da UNESP aqui mencionado e também em outras passagens refere-se à grande base textual informatizada da FCL da UNESP, Campus de Araraquara, que totaliza 80 milhões de palavras atualmente.

exemplo de *-arca*¹ e *-arca*² que a versão XXI propõe como duas formas homônimas e elementos de composição, os quais seriam usados em posição final em relação à base. E distingue *diarca, monarca, triarca* que seriam gerados de *-arca*¹ de *menarca*, gerado de *-arca*². Mas nenhuma dessas palavras foi criada no português; logo é discutível que *-arca* seja elemento de composição de nosso idioma, se o comprovante forem essas palavras. *Monarca* é um vocábulo relativamente antigo no português – é do séc. XIV – e é empréstimo do latim tardio. Quanto a *diarca*, é empréstimo erudito do alemão (1873 segundo Cunha, 1982) e *triarca* é empréstimo do inglês, que o tomou ao grego. A lição da edição de 86 é melhor. Propõe *arc(a)-* como elemento de composição para as palavras *arcebispo, arcangélico* (?), *arcipreste, arquiduque, arquimilionário*. É verdade que também *arcipreste* é empréstimo.

Contudo, não creio que existam *arquiavô* e *arquivó* que aparecem como entradas tanto no AURÉLIO como no HOUAISS (que deve ter copiado do AURÉLIO).

Ainda no AURÉLIO *fag(o)-* é corretamente interpretado como elemento de composição em *fagócito, fagocitose, fagoinibidor*, etc. Contudo, tais palavras do domínio da Biologia também podem ter entrado no português como empréstimos do francês onde elas teriam sido criadas, pelo menos as duas primeiras. Também se classifica corretamente como elemento de composição o formante *imuno-*. Ver compostos: *imunologia, imunodeficiência* (-nte), *imunodepressor, imunoenzimologia, imunopatologia, imunossupressor, imunoterápico*, etc.

Seria preciso examinar cuidadosamente neste dicionário os verbetes relativos a *elementos de composição* para distinguir o joio do trigo, ou seja, para identificar e manter aqueles que podem ser classificados assim em face da Teoria Lexical e da História do Português e aqueles que deveriam ser eliminados desse conjunto.

Já o HOUAISS derrapa violentamente nesta matéria. O dicionário está eivado de um número incrível de unidades que foram classificadas como “elemento de composição”. É tal a plethora de elementos de composição segundo este dicionário que, praticamente

em todas as suas páginas, encontramos várias ocorrências. Exemplos: *cereal(i)-, deix-, desert-, desequ-, desentros-, eleit-, envid-, epipol, fabr-, garg-, ido-, lobreg-, mend-, mendac-, nomo-, nud(i/o)-, oivar-, -olfar, plant-, pot-, prec-, put-, reboc-, red(i)-, senh-, simetr(i/o)-, temper-, trib-, etc.* Nenhum linguísta chamaria de “elementos de composição” tais unidades formais, se é que elas têm alguma autonomia. De fato, os vocábulos listados no interior de cada um destes verbetes como derivados do *soit-disant* elemento de composição são palavras cognatas.

Vejam os desses casos: *fabr-*. Nesse verbe são registrados como vocábulos formados deste elemento: *fábrica, fabricação, fabricante* [não registrado no *corpus*], *fabricando* [classificado como adjetivo!!], *fabricante, fabricável*, etc. Ignoram-se aqui dados da história da nossa língua. Aliás, algumas dessas palavras bem como outras registradas em outras dessas entradas não foram geradas dentro do português mas são empréstimos do latim, ou de outra língua latina (francês, italiano, espanhol). Na ciência linguística é muito controversa a questão das fronteiras entre a composição e a prefixação porque alguns prefixos têm uma certa autonomia – cf. *contra, extra, não, tetra, ultra*, por exemplo. É por essa razão que alguns linguístas preferem falar em *formantes*, não considerando esses operadores lexicais propriamente como prefixos. Mas essa é uma outra história. Seja como for, nenhum linguísta chamaria de “elementos de composição” as unidades formais acima arroladas.

3. A NOMENCLATURA DO DICIONÁRIO

3.1 Um *corpus* de referência

Na moderna Lexicografia qualquer obra de vulto como é o caso dos dois dicionários em exame deveria fundamentar-se em um *corpus* informatizado como fonte de referência na extração e seleção das entradas (lemas) do dicionário.

Infelizmente estes dois grandes dicionários contemporâneos do português brasileiro não seguem este modelo. Na verdade só o *Dicionário da Academia de Ciências de Lisboa*, publicado em 2001, se baseou em uma grande base informatizada de textos (obras do século XIX mas sobretudo do século XX).

O uso de um *corpus* informatizado para a confecção de dicionários implica grandes vantagens e benefícios para o produto lexicográfico, a saber:

- 1) confiabilidade dos dados como representação da língua realmente usada pelos falantes tanto em sua modalidade escrita como oral;
- 2) possibilidade de registro e identificação das fontes e de sua datação;
- 3) possibilidade de identificação dos registros ou níveis de linguagem;
- 4) abundância de dados e possibilidade de manipulação rápida desses dados;
- 5) contextualização das palavras, o que permite a extração dos valores semânticos e sintáticos das palavras com base em dados autênticos.

Operando com um *corpus* de textos e discursos efetivamente já produzidos garante-se a representatividade do acervo lexical da língua, bem como de seu uso.

3.2 Uma questão metodológica que os autores destes dicionários deveriam ter formulado de partida: qual seria a extensão da sua nomenclatura e/ou macroestrutura e qual seria o *modus operandi* para recolher este *index verborum*? De fato a seleção da nomenclatura não pode ser aleatória. Para elaborar um dicionário das dimensões do AURÉLIO (130 mil verbetes?) e do HOUAISS (230 mil verbetes?) seria preciso partir de um enorme *corpus*. Só para comparar: o *Dicionário da Língua Portuguesa Contemporânea da Academia de Ciências de Lisboa*, que tem 60.000 entradas, se baseou num *corpus* de 100 milhões de ocorrências; o *Dictionnaire de la langue française du*

XIXe. et XXe. Siècle de 70.000 verbetes usou uma base de mais de 100 milhões de palavras como fonte de recolha de dados lexicais.

O método hoje adotado é, pois: de um banco de dados informatizado faz-se a extração e a seleção da lista das palavras-entrada do dicionário. Evidentemente as palavras que ocorreram com alta e média frequência deverão integrar a nomenclatura do dicionário. Contudo, as baixas frequências (menor ou igual a cinco) no *corpus* colocam problemas. Outros critérios como a disponibilidade, a tematicidade deverão ser considerados para colher palavras no interior desse conjunto. Os *hapax legomena* (frequência 1) em princípio seriam rejeitados, pois registram idiosincrasias de autores, ou tecnicismos típicos do discurso científico muito especializado. Dentro destas categorias os dicionaristas precisam operar com muito critério, levando em conta o numerário final do dicionário planejado. Embora, na medida do possível, os dicionários gerais ambicionem dar conta do maior número de unidades do léxico, nem por isso se pode ir incluindo qualquer palavra encontrada em qualquer texto. Para que um vocábulo seja considerado como integrante do acervo social da língua, ele não pode figurar em apenas um idioleto. Assim, num grande *corpus* de textos e gêneros representativos da língua, os *hapax legomena* fazem parte das palavras que convém evitar na montagem da nomenclatura.

Nó HOUAISS a lista da nomenclatura está inchada com *criações virtuais* que não estão documentadas. Nota-se que, em vários casos, os autores registraram exemplos criados por eles próprios para validar a palavra-entrada, o que não serve como comprovante de sua pertença ao léxico da língua. Correndo o risco de ser monótona, elencarei um rol grande, mas não exaustivo, dessas criações virtuais a que faço aqui referência: *abreijerar, abrenunciação, abrenunciar, ab-renunciar, ab-renúncio, absogro, absoluto, absonar, ábsono, absorvedoiro, açafatar, acafetado, acafetar, adjeção, adormir*.² Conti-

² A citação de Fagundes Varela "adormido infante" não autoriza a criação do verbo *adormir*. Os poetas não são bons documentadores do léxico da língua porque costumam

nuando: *alcuza, aldagrante, alezirado, allfádega, alminheiro, alporquento, altibaixo, alticolúnio, arrabileiro, aspirativo* [da Língüística? – de fato é *aspirado*], *assustoso, baconista, bacorejar, dial*², *dialeta, diluviar, dobragem, doidaria, doidejante, doidejar, doidejo, embalçar, embiocado, embiocar, enfastiadiço, enfastioso, enfastismo, erronia, erubescer, esplim*.³ Continuando ainda: *espoar, faxinal, feriar, ferível, finidade, funga, fuxicaria, garreira, garrana, garrafal, gastável, gênito, gentança, gnaticídio, gronho, grosar, grossaria, imérito, imersor, imetodicamente, imiscível, imissão, imitar, imoético, impolítica, impolítico, impoluível, impulsivismo, inacusativo, incardinar, incomfortar, inesculpido, intransitado, joliz, júvida, juncada, justura, lambarão, lambarice, lâmpado, lapujice, larvejar, latifundiado, mandadeiro, mandrianar, manducação, manducativo, manducável, manita, manzona, marimacho, martelejar, etc.*

Vários desses vocábulos são latinismos tendo o(s) dicionarista (s) aportuguesado palavras latinas. Essa plethora de latinismos não documentados em textos da língua portuguesa revela um vezo de intelectuais da velha guarda. Julgam que porque é latim é vernáculo, identificando indevidamente o vocabulário do latim clássico, medieval e eclesiástico com o léxico do português. Embora o latim se constitua numa fonte permanente na qual podemos haurir e criar, a partir dele, neologismos para cunhar novos conceitos e preencher lacunas léxicas, tal eventualidade e possibilidade não autoriza o dicionarista a ir extraindo termos do léxico latino e incorporando-os ao léxico português, ou de qualquer outra língua latina.

mam evitar o vocabulário da língua comum; ademais, são muitas vezes levados a "licenças poéticas" como é exatamente esse caso.

³ Do <ingl. spleen?> – encontrado em Baudelaire, *Fleurs du Mal*, portanto na língua francesa.

4. ABRANGÊNCIA E REPRESENTATIVIDADE DO LÉXICO

4.1 A *nomenclatura* do AURÉLIO e do HOUAISS inclui um número enorme de termos técnico-científicos, regionalismos, palavras desusadas e obsoletas assim como vocábulos literários raros. Para testar a abrangência da macroestrutura desses dicionários fiz um teste. Recolhi aleatoriamente em todas as letras do alfabeto 100 palavras que ocorreram com baixa frequência (1 a 6) no *corpus* e verifiquei se elas estavam registradas nestes dicionários. Constatei que ambos os dicionários registram quase todos esses vocábulos com as seguintes exceções – *ganzé* (AURÉLIO), *cablagem* (ambos; mas registram o verbo cablar), *lacrimorréia*, *mafagafes*, *salolado*, *cabungá*; ambos, porém, registram uma variante *cabundá*. Isso mostra que ambos são muito abrangentes relativamente ao acervo lexical do português. Pode-se notar também que o HOUAISS segue de perto o AURÉLIO, o que seria de esperar já que aquele dicionário foi elaborado e publicado depois. Ademais, é evidente que o HOUAISS recolheu sua nomenclatura em todos os dicionários gerais do português que o precederam, começando com o Moraes (edição de 1813). Numerosas dentre as suas entradas são nitidamente extraídas das seguintes obras: Moraes (1813), Aulete (1888), Vieira (1871-1874) e *Cândido de Figueiredo* (1899) – tanto no que respeita à questão da datação, como relativamente ao conteúdo dos verbetes e mesmo do teor das definições.

Tanto o AURÉLIO como o HOUAISS deram primazia à *polissêmia* versus *homonímia* no processo de identificação das entradas. Isso teve conseqüências importantes no estabelecimento de suas *macroestruturas*, visto como consideraram homônimos apenas os vocábulos que têm um étimo diferente. É o caso de *acorde* (do fr. *accord*) e *acorde* (de *acordar*) seguindo o modelo do *Diccionario de la Real Academia Española* conforme informou Aurélio B. de H. Ferreira em seu prefácio e também Villar no HOUAISS. Por conseguinte, quando uma mesma forma pode nomear seja um adjetivo, seja um substantivo, ou um advérbio, os dicionaristas deram-lhe uma única

entrada, redistribuindo em subentradas as unidades léxicas que lhe são homônimas. É o caso de *andar* verbo e *andar* sub., *direito*, adj., sub., adv. Consideraram, pois, tais unidades léxicas como polissêmicas. Poderiam ter-lhes dado outro tratamento considerando-as como homônimas e assim teríamos no primeiro caso *andar*¹ verbo, *andar*², substantivo; *direito*¹, adjetivo, *direito*², advérbio e *direito*³, substantivo. Assim o critério etimológico se sobrepôs ao semântico-lexical e sintático. Esse último critério permitiria que cada categoria léxico-gramatical diferente fosse identificada como uma nova unidade do léxico. Embora tal procedimento dilate o espaço ocupado pelos verbetes, tem a grande vantagem de, em sendo lexicologicamente correto, possibilitar ao consulente, não-especialista, uma rápida identificação da palavra que procura. Além disso, muitas e muitas vezes, a mudança de categoria gramatical acarreta alterações semânticas consideráveis como é o caso de *andar* acima referido de *vista*, *visto*, etc. Há um outro inconveniente em não se separar homônimos como *andar*¹ verbo, *andar*², substantivo – a questão das lexias complexas, fraseologias e expressões que se formaram a partir dessa base. É comum que tais unidades complexas se filiem apenas a uma das duas (ou três) unidades enquadradas no mesmo verbe como ocorre neste caso e em muitíssimos outros. Assim, o dicionário terá que registrar no fim do verbe combinatórias e expressões que se reportem apenas a uma dessas categorias. No AURÉLIO, por exemplo, em *andar*: *andar naufragado*; *andar*, *virar* e *mexer*, que são derivações do verbo e não do substantivo, embora sigam a subentrada substantivo, visto como o dicionário remete tais unidades para o fim. No HOUAISS, em que a lista de lexias complexas e fraseologias registradas foi maior, no fim do verbe, encontramos: *andar aos caídos*, *andar assim*, *andar atravessado com*, *andar de mal a pior*, *andar de torto em través*, *andar fora de si*, *andar para*, *andar pingando*, *andar político*, *andar trincado*, *andar virar*, *mexer*, *a quantas anda*. Nesse exemplo como em muitíssimos outros, vê-se novamente que o HOUAISS não se fundamentou em um *corpus* de textos reais para recolher suas expressões idiomáticas. Deve tê-las colhido em glos-

sários de fraseologias e dicionários portugueses mais antigos como o *Aulete*, o *Vieira*, o *Cândido de Figueiredo*. Das duas listas do AURÉLIO e do HOUAISS só encontrei no *corpus* da UNESP: *andar de mal a pior*, *andar fora de si e a quantas anda*. Em compensação confira-se o que está documentado nesse *corpus* e que nenhum dos dois dicionários registrou: *andar a cavalo*, *andar a esmo*, *andar a galope*, *andar a paisana*, *andar a pé*, *andar a X quilômetros por hora*, *andar à frente*, *andar a jato*, *andar a passos de tartaruga*, *andar a reboque*, *andar à procura de*, *andar à toa*, *andar aos trancos (e barrancos)*, *andar às moscas*, *andar às turras*, *andar às voltas com*, *andar à vontade*, *(não) andar bem das pernas*, *andar ao léu*, *andar com as próprias pernas*, *andar de cabeça para baixo*, *andar de lá para cá (de cá para lá)*, *andar de mãos dadas*, *andar de rastros*, *andar devagar*, *andar em bando*, *andar em círculos*, *andar em ordem*, *andar em torno de*, *andar na contramão*, *andar na linha*, *andar na mira de*, *andar na moda*, *andar na ponta dos pés*, *andar nas nuvens*, *andar nos trilhos*, *andar pelos trinta [(quarenta, etc.) anos]*, *andar sem destino*, *andar sem rumo*.

4.2 Repertórios especiais

Existem alguns repertórios especiais de vocábulos que só os dicionários gerais podem abrigar e foi o que tentaram fazer estas duas obras.

Na seção "como usar o AURÉLIO Século XXI" os autores chamam de *rubrica* as *marcas de registro lingüístico*. Na verdade essas "rubricas" deveriam ser identificadas como *marcas de uso*, pois ora são marcadores da pertença da palavra-entrada a uma determinada área técnico-científica, ora indicadoras de um dado nível de linguagem, tais como: regionalismo, lusitanismo, popular, vulgar.

4.2.1 Comparando-se esta nova edição do AURÉLIO com a de 1986, constata-se que existe um número considerável de termos científicos e técnicos novos, o que é louvável. Desde a edição de 1986, já se podia notar um grande esforço para incluir termos das

mais variadas áreas como: Botânica, Biologia, Eletricidade, Estatística, Etnografia, Geologia, Genética, Marinha, Marketing, Meteorologia, Psicologia, Química, Telecomunicação, etc.

O AURÉLIO também registrou um grande número de termos das ciências matemáticas, físicas, da biologia, da botânica, assumindo claramente uma fisionomia enciclopédica. Veja-se, por exemplo, a lista de funções matemáticas: *arc cos*, *arc cosec*, *arc cosech*, etc.

4.2.2 A inclusão de *palavras raras*, pinçadas em obras literárias, remete à questão da *autoridade* literária como fonte privilegiada de coleta vocabular com fins lexicográficos. Assim, o AURÉLIO e o HOUAISS registraram muitas palavras raras que devem ter ocorrido como hapax em um texto literário, pois os artistas costumam ser amantes do inusitado. No AURÉLIO, em que estão documentadas não só as abonações dos autores mas as fontes e se pode fazer uma pesquisa reversa na versão eletrônica, encontrei muitos exemplos recolhidos em escritores brasileiros e portugueses, bem como poetas. Alguns dentre eles como Euclides da Cunha, Abelaira, Camilo Pessanha, Cornélio Pires, Coelho Neto, só para citar alguns, primam por um vocabulário raro só registrado em suas obras ou em apenas uma delas. É por causa da atribuição de autoridade literária dada ao escritor que tais palavras aparecem registradas no dicionário. Constata-se que Aurélio B. de H. Ferreira, como fizeram outros dicionaristas do passado, ficava fascinado por uma palavra exótica e a registrava em fichas para depois incorporá-la ao dicionário. Ver por exemplo no AURÉLIO vocábulos raros colhidos em poetas: *enturbar* (Raimundo Correia), *epinício* (Alberto de Oliveira), *imo* (Gonçalves Dias), *álgido*, *batalhante*, *besante e rojar* (Camilo Pessanha), *intercolúnio* (Eugênio de Castro), e em escritores: *expluir* (Raul Brandão), *acendalha*, *adinamia*, *azenegue*, *convício*, *ensiforme*, *espenda*, *exsicado*, *farragem*, *galhardear*, *intrêmulo*, *javardo*, *rechã*, *zanaga* (Euclides da Cunha), *referto* (Coelho Neto), *rabulejar*, *dé relancina*, *tacuru* (esses três em Cornélio Pires).

Diversamente, na moderna Lexicografia, considera-se que uma palavra faz parte do patrimônio léxico da língua se ela tiver sido

usada num determinado número de vezes por diferentes falantes e tiver ocorrido em mais de um tipo de texto (gênero). O AURÉLIO e o HOUAISS, porém, são típicos herdeiros da Lexicografia do passado. Segundo esse parâmetro lexicográfico, o vocabulário das grandes obras literárias deve ser registrado no dicionário como modelo ideal de língua escrita para a comunidade dos falantes por considerarem sua linguagem e seu estilo dignos de serem imitados.

Modernamente, a Lexicografia propõe uma nova atitude face ao acervo léxico da língua. O dicionário deve recolher e registrar o vocabulário em circulação em meio à comunidade dos falantes (evidentemente os mais educados, mas não apenas) documentando essa *norma lingüística* de significados e usos, que não são necessariamente literários, podendo ser, por exemplo, textos jornalísticos. Inversamente, devemos evitar as idiosincrasias dos romancistas, poetas e escritores em geral, que muitas e muitas vezes criam palavras numa situação *ad hoc*, ou por razões exclusivamente estéticas; frequentemente essas criações neológicas não se perpetuam na tradição.

4.2.3 Um problema sério diz respeito aos *regionalismos* e às palavras que o AURÉLIO classificou como *brasileirismos* desde suas duas edições anteriores (1975 e 1986). De um total de 115.243 verbetes, 24.498 entradas são de *brasileirismos* e *regionalismos*, isto é, 18% do total (Oliveira, 1999, p. 90). Na edição de 2000 repete-se o que estava nas anteriores. Dessa forma mantêm-se dados que contém informações questionáveis e que nessa última edição não foram revistos. A marca de uso *brasileirismo* só se justifica dentro de uma perspectiva do português como língua internacional por oposição às demais variedades do Português (da Europa, da África e da Ásia). Ou então para um dicionário da variedade europeia como é o caso do *Dicionário da Língua Portuguesa Contemporânea da Academia de Ciências de Lisboa* que descreve o léxico do português europeu e deve naturalmente marcar vocábulos e/ou expressões típicos do Brasil. No caso do AURÉLIO parece-me que essa marca não foi bem usada já que todo o léxico descrito é do Brasil. Quanto aos *regionalismos*, uma falha séria do AURÉLIO é não ter registrado, em

suas fontes documentais, as obras onde ele teria coletado tais *regionalismos*.

A professora Ana Maria P. P. de Oliveira (UFMS) defendeu uma tese de doutoramento na UNESP em 1999 intitulada *O Português do Brasil: Brasileirismos e Regionalismos*. Esse minucioso estudo baseado em extensa pesquisa, por mim orientada, centrou-se no dicionário AURÉLIO, extraíndo dele todos os verbetes rotulados com estas marcas sociolingüísticas. Em sua tese, a professora analisou detalhadamente os dados do AURÉLIO e tentou extrair os critérios por ele adotados para examiná-los à luz da Lingüística moderna. Constatamos que muitos senões existem neste dicionário quanto a essa matéria. Parece que mestre Aurélio usou como fontes glossários ou vocabulários de qualidade bem heterogênea. Pudemos elencar um número grande dessas obras produzidas em diversos pontos do território nacional⁴ e publicadas em várias épocas a partir do século XIX. O problema é que a qualidade científica dessas obras é questionável em muitos casos. Muitas dentre elas foram organizadas por curiosos e diletantes sem critério nenhum. Por outro lado, falta totalmente um controle objetivo para ajuizar da pertinência do registro. Por exemplo: a palavra ainda seria usada e qual a abrangência de seu uso? São vocábulos ocasionais ouvidos pelo recenseador da boca de alguém que o vocabulista registrou porque os julgou curiosos? Enfim, os problemas são muitos. Creio que Aurélio B. de H. Ferreira, amante das palavras e do vocabulário típico do Brasil, foi registrando tudo sem estabelecer parâmetros confiáveis. Assim sendo, só podemos presumir sem poder conferir. Seria necessário fazer uma gigantesca pesquisa de campo em todos os recantos do Brasil e confrontar esses dados com os *brasileirismos* registrados por Aurélio e por todos os numerosos vocabulários, glossários e dicionários que deles se ocuparam para poder dirimir dúvidas e restabelecer a verdade lexical.

Por outro lado, o HOUAISS pretende dar conta não só dos regionalismos brasileiros, mas também daqueles portugueses, bem como dos regionalismos dos países de fala lusófona. Exemplos de africanismos neste dicionário: *intentação, lamitoca, nkakana, nkanyu, nkumbi, onsenear, parcionero* (ant.); e ainda *quibala, quibeba, quibemuxi, quiçama, quiçangua*, etc. Esses últimos exemplos estão marcados como termos típicos de Angola. Não tenho conhecimento direto do trabalho dos lexicógrafos ou lexicólogos africanos, citados na introdução, como responsáveis pelas informações relativas ao léxico africano. Assim neste particular não posso ajuizar. Sobre esta matéria, na introdução do HOUAISS [cf. 10.h] afirma-se: “por vezes, não foi possível localizar geograficamente determinado provincialismo português ou, mais raramente, um regionalismo brasileiro”. Como podem estes dicionaristas ter tal certeza sem realizar pesquisa? No que concerne os regionalismos do Brasil parece evidente que o HOUAISS usou o AURÉLIO como fonte de informações, que tem os problemas já referidos. E quais os especialistas portugueses que respondem pelos regionalismos de Portugal?

4.2.4 Em ambos os dicionários há muitas palavras *obsoletas* colhidas em dicionários do passado, o que deveria ter sido assinalado. Cf. *almofia, almofreixe, almogaravia, cavilar, ceno, cenosidade, cerome, deporte, encruelecer, esbraguilhado, fuzilhão, joldra, justura, nadível, nadivo, perecedoiro, perfazimento, perilha, sufragâneo, sumilhar, sumpto, tabardo, tábido, tacanheza, tacanhice*.

4.2.5 Foram também incluídas as denominações dos *gentílicos* de todos os municípios do Brasil, relação essa que poderia ter sido alistada num anexo, dado seu interesse menor para um dicionário de língua. Conferir: *aliancense* [natural de Aliança, PE], *andaraiense* [natural de Andaraí, BA], *areia-branquense* [natural de Areia Branca, RN], *fabricianense* [natural de Coronel Fabriciano, MG], *jacareiense* [natural de Jacaré, SP], *mambucabense* [natural de Mambucaba, RJ], *raul-soarense* [natural de de Raul Soares, MG], *reboucense* [natural de Rebouças, PR], *recreense* [natural de de Recreio, MG], *sousense* [natural de Sousa, PB], etc. Ver também a seqüência de *gentílicos* forma-

dos com *porto*: *porto-alegrense, porto-amazonense, porto-belano, porto-calvense, porto-felicense, porto-firmense, porto-folhense, porto-franquino, porto-lucenense, porto-pedrense, porto-realense, porto-riquenho, porto-velhense*. E o AURÉLIO Séc. XXI acrescentou ainda: *porto-alegrense-do tocantins, porto-belense, porto-estrelense, porto-gandrense, porto-gauchense, porto-miense, porto-mozense, porto-segurense, porto-unionense, porto-veracruzense, porto-vitoriense, porto-waltense, porto-xavierense*. Foi um expediente um tanto inútil para ampliar o número de verbetes do dicionário.

5. ESTRUTURA E TRATAMENTO DOS VERBETES

5.1 Encontramos *definições*, às vezes, erradas. Cf. *Lexeologia* – termo raro, não usado na terminologia lingüística. O AURÉLIO remete a *Lexicologia*. Entretanto, a definição de *Lexicologia* não é adequada; de fato trata-se de uma das ciências autônomas do léxico e não, parte da gramática, como aí se diz. É comum tanto no AURÉLIO, mas sobretudo no HOUAISS, a palavra-entrada ser definida com vocábulos incomuns e até raros. Seria desejável que as definições fossem feitas com um vocabulário usual, e preferivelmente, um vocabulário básico de umas duas mil palavras mais frequentes.

5.2 Com alguns pequenos senões muitas vezes a *estrutura do verbete* é melhor no AURÉLIO do que no HOUAISS. Cf. *nervoso*, p.ex. Nesse caso é criticável, apenas, o uso da palavra *nerval* no AURÉLIO para definir, pois essa palavra, se usada, é rara, visto como não ocorreu no *corpus* da UNESP. As *lexias* complexas (*fibra nervosa, tique nervoso*) e unidades fraseológicas complexas (*fibra nervosa amielínica, fibra nervosa mielínica, sistema nervoso autônomo, sistema nervoso central*) das terminologias científicas aparecem no fim do verbete, de modo destacado e com a respectiva marca de domínio científico. Especialmente na versão eletrônica, fica claro para o consulente cada um desses detalhes enfatizados pelo uso de cores distintas.

Cf. ainda os seguintes exemplos ilustrativos tirados do AURÉLIO:

cabeçalho [De cabeça + -alho.] S. m.

1. Timão do carro, do qual pende a canga.
2. Título de jornal ou de outra publicação periódica, que compreende data, número, periodicidade, etc.; cabeço.
3. Título destacado de artigo, notícia, etc.
4. Título de capítulo.
5. Dizeres que encimam as colunas e casas de uma tabela, as páginas de um livro em branco, ou certos formulários e fichas. [Cf., nesta acepç., entrada (22).]

A 1ª acepção é hoje rara; não devia, pois, constar em primeiro lugar. As acepções 2., 3. e 4. poderiam ser sintetizadas em apenas um sentido e serem alçadas à posição de primeira acepção por ser o sentido mais comum.

E também:

cereal [Do lat. cereale, 'referente a Ceres, a deusa das searas'.]

Adj. 2 g.

1. Diz-se das gramíneas e doutras plantas (trigo, cevada, milho, etc.) cujas sementes, transformadas em farinha, servem para a alimentação.
 2. Relativo a pão.
- S. m.
3. Planta cereal (1).
 4. Searas, messes.

[Como s. m., é m. us. no pl. Cf. cirial.]

Esta definição não está boa. Aí se dá como significado básico, o do adjetivo. Ora, em 129 ocorrências (sing. + pl.) do nosso *corpus* nem uma só delas era um adjetivo; todas eram substantivo. Vejam-se os exemplos abaixo, extraídos do nosso *corpus* e que ilustram o valor semântico e o uso desta palavra:

Os cereais eram plantados visando principalmente à subsistência.

Para o vice-presidente da Bolsa de Cereais de São Paulo...

Nos silos, onde o cereal é estocado a granel...

OAT BRAN (Farelo de Aveia) é um cereal 100% natural. (numa propaganda)

Já o verbete *ciclo* está bem definido e formulado. O AURÉLIO parte dos sentidos mais gerais para os mais específicos, alistando, na seqüência, os significados técnicos das diversas áreas científicas tais como: *ciclo anovulatório* (fis.), *ciclo biológico* (biol.), *ciclo de Chandler* (astr.), *ciclo de Krebs* (bioquim.), etc.

Outros exemplos de definição geralmente boa com pequenos senões: *contínuo*, *perene*, *permanente*. O método de integrar o substantivo ao adjetivo numa entrada única em função do critério adotado em relação aos homônimos e dando primazia ao adjetivo como se vê nesses dois verbetes *contínuo* e *permanente* não foi feliz. Embora não seja o caso desses dois vocábulos, muitas e muitas vezes o substantivo é mais usual que o adjetivo; outras vezes, o valor semântico individualiza muito o substantivo como é o caso tanto de *contínuo* como de *permanente*. Assim, como já se disse, a melhor metodologia seria considerar-se esses vocábulos como um caso de homonímia e separar sempre a categoria adjetivo da categoria substantivo, o que é legítimo, lingüísticamente falando.

Usarei alguns verbetes mais para comentar alguns aspectos e problemas que reiteradamente se verificam no dicionário AURÉLIO. Começemos com o verbete *livre*.

O dicionarista inclui *sentidos* que já não se usam, caso tenham sido de fato usuais em etapa anterior da língua. Julgo que frequentemente induzido por registros de dicionários que o precederam como o AULETE, por exemplo (1ª ed. de 1881?) inclui sentidos inexistentes na língua contemporânea. É o caso das acepções 14 [*livre* = desregrado, licenciado], 15 [espontâneo, natural], 16 [que não tem limites: imenso, infinito]. O HOUAISS faz o mesmo.

Quanto às *combinatórias freqüentes* que podem gerar, ou não, novas unidades lexicais, o AURÉLIO registra algumas; contudo, muitas vezes registra lexias que são raras, pois não ocorreram no *corpus* da UNESP e suspeito que são dados pouco confiáveis copiados de outros dicionários. Cf. por exemplo: *atmosfera*~, *bem (bens) livre(s)*; *cálice*~; *livre-culto, livre-cultista, pulso* ~; *tiro* ~.

Não encontrei nenhuma ocorrência de *livre-câmbio* mas, sim, *câmbio livre* que, aliás, é a lexia complexa (termo) registrada em dicionários de economia.

Algumas combinatórias lexicais parecem ser cópias do AULETE tais como *vida livre* com significado de <desregrado, licenciado>. Quanto à *livre vontade*, acho que é um equívoco; deveria constar, sim, a expressão estereotipada “de livre e espontânea vontade”. Outro decalque do AULETE: *livre-alvedrio*. Fazendo novamente o confronto com o nosso *corpus* constatei não haver nenhuma ocorrência de *livre-alvedrio*. Registram-se 4 ocorrências de *alvedrio* (palavra isolada) nesse *corpus*.

Na listagem de lexias complexas feitas ao fim do verbete algumas são combinatórias freqüentes com *livre* e deveriam ser referidas como tal. Faltam aqui: *imprensa livre, queda livre*. O AURÉLIO inclui corretamente como entradas autônomas: *livre-arbítrio, livre-câmbio* (melhor *câmbio livre*), *livre-comércio, livre-docência, livre-docente, livre-iniciativa, livre-pensador, livre-troca*. Contesto a existência como unidades do léxico de: *livre-cambismo, livre-cambista, livre-cultismo, livre-cultista, livre-culto, livre troca*. Por outro lado, o dicionário deveria ter registrado como combinatórias freqüentes e/ou lexias complexas: *amor livre, boca-livre, livre circulação (de mercadorias, capitais), livre comércio, livre competição, livre concorrência, livre debate, livre-empresa, (direito de) livre escolha, livre flutuação de (dos) preços, livre iniciativa, livre negociação, livre trânsito*.

Algumas dessas lexias complexas foram adequadamente registradas em verbetes que têm como lema um substantivo com que *livre* se combina: *feira livre em feira, vão livre em vão, mercado livre em mercado* [embora também ocorra com freqüência *economia de livre*

mercado, sistema de livre mercado com anteposição] e livre curso como fraseologia verbal dar livre curso a em curso.

Ainda com respeito à estrutura do verbete. Vou comentar outro verbete em que não houve muito critério de ordenação dos dados lexicais – o verbete *direito*. O primeiro problema é o já referido da não separação dos homônimos *direito* adj., *direito* subst. e *direito* adv. Mas não é essa a questão que nos ocupa agora. A subentrada substantivo arrola, ao fim do verbete, uma grande lista muito heterogênea de combinatórias, a saber: *direito adjetivo, ~ administrativo, ~ adquirido, ~ aéreo, ~ agrário, ~ assistencial, ~ autoral, ~ cambiário, ~ canônico, ~ civil, ~ clássico, ~ comercial, ~ constitucional, ~ consuetudinário, ~ costumeiro, ~ criminal, ~ das gentes, ~ de arena, ~ de fundo, ~ de petição, ~ de preferência, ~ de regresso, ~ de reprodução, ~ de resposta, ~ de retorno, ~ do trabalho, ~ escrito, ~ falencial, ~ falimentar, ~ financeiro, ~ fiscal, ~ individual, ~ industrial, ~ internacional privado, ~ internacional público, ~ intertemporal, ~ judiciário, ~ líquido e certo, ~ marítimo, ~ natural, ~ normativo, ~ objetivo, ~ penal, ~ personalíssimo, ~ pessoal, ~ político, ~ positivo, ~ privado, ~ processual, ~ público, ~ real, ~ regressivo de recurso, ~ romano, ~s conexos, ~s de estola, ~s de mercê, ~s de pé-de-altar, ~ subjetivo, ~ substantivo, ~ tributário*. Ora, nesse rol temos: lexias compostas que são termos da ciência do direito: *direito administrativo, direito agrário, direito autoral, direito canônico, direito civil, direito comercial, direito constitucional, direito do trabalho, direito fiscal, direito industrial, direito internacional privado, direito internacional público, direito penal, direito processual, direito público, direito romano, direito tributário*. E ainda: sintagmas freqüentes na linguagem especializada do direito e outros que caíram em desuso. O dicionarista deveria ter distinguido alhos de bugalhos.

5.2 Quando a palavra tem vários *sentidos*, um problema muito complicado é isolar os principais sentidos sem se deixar confundir pelas conotações resultantes do contexto; e dessa forma individualizar apenas os significados básicos em que, de fato, se podem reconhecer semas diferentes. Creio que, muitas vezes, essa constitui

uma fragilidade tanto do AURÉLIO como do HOUAISS. Ao invés de tentar fazer um esforço de separação de sentidos e de sua ordenação, os dicionaristas vão expandindo o verbete e acrescentando mais e mais acepções, quando o ideal seria sintetizar e congregar várias delas numa mesma acepção.

5.3 O HOUAISS incluiu na estrutura do verbete um espaço para as redes de significação do léxico buscando inovar, não só em relação ao AURÉLIO, mas também em relação à Lexicografia tradicional em Língua Portuguesa. Na nossa tradição só dicionários especiais se ocupam de modo sistemático de sinônimos e antônimos. Embora o propósito fosse louvável, não me parece que o HOUAISS foi feliz em sua inovação, dada a forma como tratou a controversa questão da sinonímia e da antonímia. Pretendendo ser exaustivo, acaba por incluir como sinônimos palavras que não o são. De fato, como sabemos, os sinônimos verdadeiros, totais, são raros; assim, o dicionarista deveria ter informado que indicaria não só os sinônimos, mas também os parassinônimos. E o mesmo se diga dos antônimos e dos contrários. Como não tinha uma visão teórica bem estabelecida sobre essa matéria, o HOUAISS extrapola de modo assustador. No exemplo dado a seguir do verbete *permanente*, o dicionário arrola como sinônimos desta palavra: *arraigado, arreigado, constante, contínuo, definitivo, duradouro, durável, efetivo, estável, eternal, eterno, eviterno, fiel, firme, fixo, imanente, imorredouro, imortal, imperecível, imperturbável, imudável, imutável, inabalável, inalterável, incessante, indelével, indubitável, inexorável, infindável, infinito, ininterrupto, inquebrantável, interminável, invariável, inviolável, irrevocável, irrevogável, leal, perdurável, perenal, perene, permanecente, perpétuo, perseverante, persistente, seguro, sempiterno, sólido, tenaz, vago*. E como antônimos: *alterável, breve, caduco, descontinuado, descontínuo, efêmero, extingüível, findável, finito, frágil, fugaz, fugidio, fugitivo, impermanente, incerto, inconcluso, inconstante, indefinido, infreqüente, inseguro, instável, intercadente, interino, intermitente, interrompido, interrupto, intervalado, lábil, ligeiro, momentâneo, morredouro, mortal, movediço, móvel, mudadiço, mudável, mutável, ocasional, passadiço, passageiro, perecedor, perecedouro,*

perecível, perituro, perturbável, precário, provisional, provisório, temporário, terminável, transitivo, transitório, variável, vário, versátil, voante, volátil, voltário, voltívolo, volúvel. É evidente que um grande número desses adjetivos podem ser considerados, no máximo, como integrantes do campo semântico-lexical de *permanente*, mas não como sinônimos dessa palavra e nem como antônimos. Aliás, muitos deles fazem parte daquelas criações lexicais já referidas, inventadas pelo HOUAISS. Não fiz um confronto minucioso, mas creio que o HOUAISS se serviu de uma fonte problemática para colher sua pletora de sinônimos e antônimos – o dicionário analógico de Azevedo (1974?), que está arrolado na bibliografia deste artigo. Embora o dicionário de Azevedo possa prestar serviços a um lexicógrafo e a um escritor, ele tem que ser usado com grande cautela, pois está eivado de “alucinações” semântico-lexicais. Em suma, não só no exemplo dado, mas em muitos outros, o consulente precisa duvidar de muitas informações registradas pelo HOUAISS no capítulo relativo a sinônimos e antônimos. Um dicionário de ótima qualidade como o *Petit Robert*, econômico em relação a sinônimos e antônimos, teria sido um bom modelo para o HOUAISS.

6. ETIMOLOGIA E HISTÓRIA DAS PALAVRAS

A etimologia e a datação das palavras que se anunciou como uma marca de superioridade do dicionário HOUAISS sobre os dicionários contemporâneos também é questionável. A língua portuguesa não possui estudos confiáveis sobre a história de seu léxico em seu conjunto para que se possam fazer afirmações categóricas para um número considerável de palavras do nosso vocabulário. Com exceção dos dicionários etimológicos/históricos de Antenor Nascentes, José Pedro Machado e Antônio Geraldo da Cunha, não possuímos muitas fontes para pesquisar este tipo de informação. E essas obras não são isentas de incorreções, além de não cobrirem toda a imensidão do léxico. Em um número significativo de casos suas lições

têm que ser revistas e reformuladas. Portanto, não creio que se possa emprestar fidedignidade a muitas informações de cunho etimológico e histórico propostas pelo HOUAISS.

Nas primeiras páginas do dicionário HOUAISS são elencadas suas fontes de referência, que impressionam pelo volume e pelo que revelam de erudição. Contudo, não creio que os autores deste dicionário tenham lido aquelas obras e as consultado. Esporadicamente pode ser. Os responsáveis pela etimologia que talvez tenham sido também responsáveis pela história de muitas palavras, cuja evolução semântica se tentou elaborar, não têm um curriculum científico no domínio da Linguística Histórica ou da Filologia Portuguesa para constituírem autoridade sobre os ensinamentos que passam aos consulentes. Examinando-se as fontes de datação e etimologia tem-se a impressão de que estamos diante da lista das obras da biblioteca de Houaiss. Por outro lado, várias obras aí alistadas têm interesse para um dicionarista e um intelectual interessado na história das línguas como ele era, mas não se constituem propriamente numa fonte de datação e de etimologia para o vocabulário do português. Dentre outros, alguns dicionários históricos e etimológicos, por exemplo: *The Century dictionary. An Encyclopedic lexicon of the English language* (1889-1891); *Dictionnaire étymologique de l'ancien français* (1992); *Dictionnaire historique de la langue française* (1974); *Trésor de la langue française*; *The Oxford dictionary of English etymology*; e vários outros de línguas eslavas, do alemão, etc. Também se incluem duas edições do dicionário *Webster* – 1961 e 1996. Se a obra *Portugaliae Monumenta Historica* é fonte importante para a história do português, não me parece que dicionários etimológicos do inglês, do russo, e mesmo do francês possam sê-lo. Não ajudam nem a fazer a história das palavras na evolução diacrônica do português, nem podem ser considerados como fontes de datação, pois trata-se da história de outras línguas.

Foi uma temeridade o mentor do dicionário se ter lançado a um empreendimento tão ambicioso como este sem contar com uma equipe sólida de filólogos que respondesse por uma área tão comple-

xa e que exige a dedicação de uma vida, visto como é tão pequena a produção científica sobre a história do português e sobretudo de seu vocabulário. Julgo que filólogos de sólida formação científica teriam provavelmente desencorajado Houaiss de se lançar a uma tal façanha científica sem ter os recursos para isso.

O modelo de verbete do HOUAISS foi estruturado tendo como eixo as datações de registro escrito da palavra, bem como de seus diferentes sentidos, seguindo o modelo do grande *Oxford English Dictionary*. Como já se afirmou essa foi uma opção temerária, pois a língua portuguesa não dispõe de muitas fontes confiáveis que permitam ao lexicógrafo estabelecer tais datações. O HOUAISS valeu-se sobretudo das obras e pesquisas de Antônio Geraldo da Cunha para esse fim. Além de problemático o registro do primeiro uso da palavra, isso tem conseqüências na estruturação do verbete. Assim, em princípio, o HOUAISS ordena as acepções em ordem cronológica, assumindo, pois, uma postura de dicionário histórico e não de um dicionário de usos da língua contemporânea. Convém lembrar que o Oxford, que lhe serviu de modelo, pelo contrário, se apresenta como um dicionário baseado em princípios históricos. Por conseguinte, um consulente comum do HOUAISS, desconhecedor da diacronia da língua, ficará surpreso ao constatar que o sentido mais comum na atualidade pode aparecer em 2º, 3º, ..., 8º lugar na sucessão histórico-hierárquica dos sentidos. Vejam-se três exemplos para ilustrar.

Botão. A primeira acepção indicada é <gema; a flor antes de desabrochar>. O sentido <pequeno objeto para fechar uma veste (blusa, camisa, etc.)> aparece como 8ª acepção. Consultando-se o *Diccionario Crítico Etimológico de la Lengua Castellana* de J. Corominas, que é uma das boas fontes para o estudo da história lexical das línguas latinas da Ibéria, vê-se que o sentido <pequeno objeto para fechar uma veste (blusa, camisa, etc.)> foi o primeiro significado registrado para o espanhol – século XIII. É improvável que, dada a intensa convivência e interação entre o reino de Portugal e os reinos castelhano:; nessa época, tal significado fosse exclusivo do es-

BIDERMAN, Maria Tereza Camargo. *Análise de dois dicionários gerais do Português Brasileiro contemporâneo: o Aurélio e o Houaiss*.

panhol. Baseando-se em Cunha, o HOUAISS atribui sua primeira acepção ao século XIV, precedendo aquele oitavo sentido. É preciso pesquisar mais para se estabelecer tal cronologia.

Faxina. A acepção <conjunto de gravetos, lenha miúda> que foi usada no passado [cf. Morais, 1813] hoje é praticamente desconhecida. É o primeiro sentido registrado neste verbete.

Surgir. A acepção <lançar âncora ou amarra no porto e dar fundo com elas; ferrear, ancorar, fundear> aparece como primeira acepção, sentido esse que já caiu em desuso. Também o sentido mais comum de *surgir* no português contemporâneo <aparecer> só é registrado como 8ª acepção.

É evidente que a ordenação dos sentidos das palavras no verbete causará espanto e até dificuldades para os consulentes.

7. CONCLUSÃO

Numa avaliação geral, é preciso admitir que o AURÉLIO continua sendo um dicionário mais coerente e de melhor qualidade técnica que o HOUAISS, apesar das muitas críticas feitas neste artigo. Embora se tenha afirmado que o recurso a obras literárias como fonte e aval autorizado de uso não seja indispensável, ainda assim essa é uma qualidade que o AURÉLIO possui indubitavelmente e o HOUAISS, não. E bem visível na versão eletrônica de fácil consulta. Quando o consulente constata que um grande escritor usou aquela palavra e pode conferir, de imediato, o enunciado em que ela ocorreu, sentir-se-á mais seguro em usá-la. É verdade que o dicionário AURÉLIO poderia ser de melhor qualidade, corrigindo os desacerdos do dicionarista Aurélio em suas versões anteriores (1975 e 1986) e aprimorando recursos que já possuía e, sobretudo atualizando a nomenclatura, incluindo neologismos entrados para o idioma depois destas datas.

Infelizmente, é preciso cobrar das equipes que elaboraram estas obras, o fato de um trabalho tão grande e mesmo insano ter

Filol. lingüist. port., n. 5, p. 85-116, 2002.

falhas que poderiam ter sido evitadas. De fato, se um corpo de linguistas de formação especializada em Lexicologia/Lexicografia/Terminologia tivesse cooperado nestes grandes empreendimentos e se essas obras se tivessem fundamentado em um grande *corpus* informatizado, representativo do Português Brasileiro contemporâneo, elas poderiam atingir o status de grande tesouro lexical da nossa língua.

BIBLIOGRAFIA

- ALVES, I. M. (1990) *Neologismo. Criação lexical*. São Paulo, Ática.
- AULETE, F. J. C. (1881) *Diccionario contemporaneo da lingua portugueza*. Lisboa, Parceria Antonio Maria Pereira, 2 vols.
- AZEVEDO, F. F. dos S. (1974?) *Dicionário analógico da Língua Portuguesa*. Brasília, Coordenada – Editora de Brasília.
- AZEVEDO, G. de (1966) *Vocabulário do criatório norte-riograndense*. (Estudos Brasileiros n. 20). Rio de Janeiro, Edições Siá.
- BASILIO, M. (1987) *Teoria lexical. Série Princípios*. São Paulo, Ática.
- BEAUREPAIRE-ROHAN (1956) *Dicionário de vocábulos brasileiros*. 2. ed. Salvador, Livraria Progresso Editora.
- BIDERMAN, M. T. C. (2001) *Teoria lingüística. Teoria lexical e lingüística computacional*. São Paulo, Martins Fontes. Nova edição atualizada.
- _____. (1985) *Lexicografia e Lexicologia* (organizadora e autora de: introdução, glossário e "A ciência da Lexicografia" e "O dicionário padrão da língua"). Suplemento Especial da Revista ALFA. São Paulo, EDUNESP.
- _____. (1998) *O estado da arte nas ciências do léxico: lexicologia, lexicografia e terminologia*. ALFA. v. 42 (n. esp.). São Paulo, EDUNESP. (organizadora e autora).
- _____. (1998) "Estrutura estatística do léxico do português brasileiro contemporâneo". *Actes du XXIIe. Congrès International de Linguistique et de Philologie Romanes*. Bruxelles, 23-29 jul. 1998. Vol. IV Des mots aux dictionnaires, p. 93-102.
- _____. (1999) Conceito lingüístico de palavra. In BASÍLIO, M. *Palavra. Série Linguagem*. Volume temático I. PUC. Rio de Janeiro, Grypho, p. 81-97.
- BLUTEAU, D. R. (1712-28) *Vocabulário português e latino*. Coimbra e Lisboa, várias impressoras. 8 v. e 2 supl.
- BOLÉO, M. de P. (1943) *Brasileirismos: problemas de método*. Coimbra, Coimbra Editora. [Separata de Brasília, vol. III].
- CARDOSO, A. L. (1961) *Amerigenismos*. Rio de Janeiro, Biblioteca do Exército Editora.

BIDERMAN, Maria Tereza Camargo. Análise de dois dicionários gerais do Português Brasileiro contemporâneo: o Aurélio e o Houaiss.

CLEROT, L. F. R. (1959) *Vocabulário de termos populares e gírias da Paraíba*. 1. ed. Rio de Janeiro.

CORAZZARI, O. (1992). *Phraseological units*, Consiglio Nazionale delle Ricerche. Istituto di Linguistica Computazionale. Network of European Reference Corpora (NERC). serial n. 68. Pisa (manuscrito).

CORBIN, D. (1987) *Morphologie dérivationnelle et structuration du lexique*. Tübingen, Niemeyer, Linguistique Arbeiten. 2 vol.

COROMINAS, J. (1954) *Diccionario crítico etimológico de la lengua castellana*. Berna, Ed. Francke.

COSTA, F. A. P. (1937) *Vocabulário Pernambucano*. Separata do vol. XXXIV da Revista do Instituto Archeológico Histórico e Geográfico Pernambucano. Recife, Imprensa Official.

CUNHA, A. G. (1982) *Dicionário Etimológico da Língua Portuguesa*. Rio de Janeiro, Nova Fronteira.

D'ALBUQUERQUE, A. T. (s.d.) *Falsos brasileirismos*. Rio de Janeiro, Getúlio Costa.

Dicionário da Língua Portuguesa Contemporânea. Academia de Ciências de Lisboa (2001). Lisboa, Editorial Verbo. 2 vol.

Dicionário Eletrônico Aurélio (2001) Versão 3.0. Rio de Janeiro, Nova Fronteira.

FERREIRA, A. B. de H. (1975) *Novo Dicionário da Língua Portuguesa*. Rio de Janeiro, Nova Fronteira.

_____. (1986) *Novo Dicionário da Língua Portuguesa*. Rio de Janeiro, Nova Fronteira.

GARCIA, R. (1915) *Diccionario de brasileirismos: peculiaridades pernambucanas*. Separata da Revista do Instituto Histórico e Geográfico Brasileiro, v. 76, n. 1, p. 637-947, Rio de Janeiro.

GROSS, M. (1982) Une classification des phrases figées du français. *Revue québécoise de linguistique*. v. 11, n. 2, p. 151-85.

HAENSCH, G.; WOLF, L.; ETTINGER, S.; WERNER, R. (1982) *La Lexicografía*. De la lingüística teórica a la lexicografía práctica. Madrid, Editorial Gredos.

HOUAISS, A.; VILLAR, M. de S. (2001) *Dicionário Houaiss da Língua Portuguesa*. Rio de Janeiro, Objetiva.

LARA, L. F. (1992) El discurso del diccionario. In WOTJAK, G. (ed.) *Estudios de Lexicología y Metalexigrafía del Español Actual*. Tübingen, Max Niemeyer Verlag, p. 1-12.

LIMA, H.; BARROSO, G. (orgs.) (1939) *Pequeno Dicionário Brasileiro da Língua Portuguesa*. 2. ed. Rio de Janeiro/São Paulo, Civilização Brasileira.

MARROQUIM, M. (1934) *A língua do Nordeste* (Alagoas e Pernambuco). São Paulo, Companhia Editora Nacional.

MENDONÇA, R. (1935) *A influência africana no português do Brasil*. São Paulo, Companhia Editora Nacional.

MEYER-LÜBKE, W. (1935) *Romanisches etymologisches Wörterbuch*. Dritte neu bearbeitete Auflage. Heidelberg.

MIRANDA, V. C. (1968) *Glossário Paraense*. Universidade Federal do Pará. 1. ed. 1905. Coleção de Vocábulos Peculiares à Amazônia e especialmente à Ilha de Marajó.

MORAIS SILVA, A. de (1813) *Diccionario da lingua portugueza*. 2. ed. Lisboa, 2 vol.

Filol. lingüíst. port., n. 5, p. 85-116, 2002.

MURAKAWA, C. de A. A. (1984) *O primeiro dicionário de Língua Portuguesa de Antonio Morais Silva*. Estudo crítico da edição de 1813. Dissertação de mestrado. Faculdade de Ciências e Letras, UNESP, Araraquara.

NASCENTES, A. (1967) *Dicionário da Língua Portuguesa*. Rio de Janeiro, Dep. de Imprensa Nacional. 4 v.

Nova Aurélio (1999) *O Dicionário da Língua Portuguesa*. Século XXI. Rio de Janeiro, Nova Fronteira.

NUNES, Z. C. e NUNES, R. C. (1996) *Dicionário de regionalismos do Rio Grande do Sul*. 7. ed. Porto Alegre, Martins Livreiro Editora.

OLIVEIRA, S. A. (1940) *Expressões do populário sertanejo*. São Paulo, Instituto Histórico e Geográfico.

OLIVEIRA, A. M. P. de (1999) *O Português do Brasil: brasileirismos e regionalismos*. Tese de Doutorado. Faculdade de Ciências e Letras, UNESP, Araraquara.

_____.; ISQUERDO, A. N. (org.) (1998) *As Ciências do Léxico: Lexicologia, Lexicografia e Terminologia*. Campo Grande, Editora da Universidade Federal do Mato Grosso do Sul, p. 129-42.

ORTÊNCIO, W. B. (1983) *Dicionário do Brasil Central*. Subsídios à filologia. São Paulo, Ática.

PICOCHÉ, J. (1992) *Précis de lexicologie française*. Paris, Nathan.

REY-DEBOVE, J. (1984) Léxico e dicionário. *Alfa* (suplemento). Centro de Publicações da UNESP, São Paulo, p. 45-69.

RIO-TORTO, G. (1998) *Morfologia derivacional. Teoria e aplicação ao Português*. Porto, Porto Editora.

ROBERT, P. (1990) *Le Petit Robert*. Paris, Les Dictionnaires-Robert.

SINCLAIR, J. M. et alii. (1987) *Looking up*. An account of the COBUILD project in lexical computing. London, Collins Cobuild.

SERAINE, F. (1959) *Dicionário de termos populares*. Rio de Janeiro, Organizações Simões.

SOARES, A. J. Macedo (1954-1955) *Dicionário Brasileiro da Língua Portuguesa*. 2 vol. São Paulo, Revista dos Tribunais/Instituto Nacional do Livro.

VIEIRA, F. D. (1871-1874) *Grande Diccionario Portuguez ou thesouro da Língua Portuguesa*. Porto, Editores Chardron e Bartholomeu H. de Moraes, Rio de Janeiro.

Webster's New Twentieth Century Dictionary of the English Language Unabridged (1972) 2. ed. Simon and Schuster, New York.

BIDERMAN, Maria Tereza Camargo. Análise de dois dicionários gerais do Português Brasileiro contemporâneo: o Aurélio e o Houaiss.

ABSTRACT: This article makes an analysis of two general dictionaries recently published: *Dicionário HOUAISS da Língua Portuguesa* (2001) and the last edition (2000) of *AURÉLIO Século XXI* dictionary. Primarily I examined the theoretical basis of these works and the flaws concerning concepts and methods of Lexical Theory; I also indicate the need for dictionaries to be based on texts that were actually printed or produced. The wide-ranging and representativity of lexicon registered in these dictionaries and the criteria for choosing word-entries are analysed, as well as the treatment applied to the articles of these dictionaries.

Keywords: dictionary, general dictionary, lexicography, computerized *corpus*, macrostructure and microstructure of the dictionary.