

GLOSSÁRIO

Maria Tereza Camargo BIDERMAN*

abonação

Frase ou enunciado, extraído de um bom autor da língua, onde ocorre a palavra que está sendo definida e/ou explicada no dicionário. Ilustra o uso da palavra no contexto. Cf. documentação.

acepção

Cada um dos sentidos ou significados de uma palavra polissêmica, isto é, que tem vários valores semânticos. Cf. polissemia.

algoritmo

Conjunto de regras operatórias, ou seqüência de raciocínios e/ou operações que podem fornecer a solução de um problema. Em Linguística Computacional chama-se de *algoritmo* um processo de cálculo que associa de uma maneira unívoca um conjunto de dados lingüísticos a um conjunto de resultados. Esse processamento se faz por um conjunto finito de operações e de regras explícitas e obrigatórias, que devem ser compreendidas e executadas numa determinada ordem por um computador. Um *algoritmo* pode, por exemplo, executar automaticamente a decomposição morfológica de um enunciado. Cf. analisador sintático e semântico.

analisador sintático (parser) e semântico

Um *analisador sintático e semântico* é constituído de um conjunto de algoritmos, elaborados e ordenados de tal forma que um computador possa analisar qualquer frase ou enunciado que se lhe apresente, desde que a máquina siga as instruções do analisador. Em princípio, um *analisador sintático e semântico* foi criado para analisar uma determinada língua. Como as línguas são diferentes estruturalmente, faz-se necessário criar um analisador para cada língua, ou adaptar um modelo que se tenha revelado eficiente em outra língua.

antônimo

Palavra que se opõe a uma outra pelo sentido. Ex.: *frio* é antônimo de *quente*; *bonito*, de *feio*.

arquilexema

Palavra superordenada semanticamente em relação a outras palavras porque o seu significado é mais geral; assim, pode aplicar-se a vocábulos mais específicos, podendo ser usado como um substantivo para esses últimos. Ex.: as palavras *coisa*, *objeto* são arquilexemas; não contêm nenhum sema (unidade de significação) específico. Podem, portanto, substituir palavras como: *cadeira*, *banco*, *automóvel*. Nesse tipo de unidade

* Departamento de Linguística — Instituto de Letras, Ciências Sociais e Educação — UNESP — 14800 — Araraquara — SP.

léxica (arquilexema) estão neutralizados todos os traços específicos dos lexemas individualizantes como *cadeira, banco, automóvel*.

arquivo

Conjunto de dados informativos armazenados, depois de terem sido classificados e ordenados, a fim de serem utilizados posteriormente. Em Lexicografia é o conjunto dos dados léxicos que servirão para a redação de um dicionário. A documentação sobre esses dados, classificados em fichários, tem, como unidade de informação, a ficha. Em Informática é qualquer conjunto de dados (banco de dados), classificados e ordenados para poderem ser recuperados facilmente quando se fizerem necessários. Em Linguística Computacional tem o mesmo significado; nesse caso, os dados do arquivo ou banco de dados são relativos à língua, ou a várias línguas.

binário

Composto de duas unidades ou dois elementos. Em matemática um sistema binário é aquele que opera apenas com duas unidades nas suas regras operatórias. O computador é uma máquina que opera na base binária.

binarismo

Sistema baseado no princípio da oposição binária. Muitos lingüistas reconhecem nas línguas a presença e atuação deste princípio. Assim, encontramos em fonologia claras evidências de oposição binária entre os fonemas. No português, por exemplo, surdos (/p/, /t/, /k/) versus sonoros (/b/, /d/, /g/). No léxico também esse princípio parece atuar na estruturação dos significados, mas de maneira menos sistemática. Por exemplo: *homem X mulher; frio X quente; menino X menina etc., campo léxico ou lexical, campo semântico*.

Conjunto de unidades léxicas que se referem mutuamente e cujos valores semânticos se delimitam reciprocamente. Pode-se compor tanto de significantes derivados de uma mesma base formal (*estudo, estudante, estudar, estudioso*) como dos significados das palavras de valor semântico próximo (*trabalho, labor, tarefa, carga, labuta*). Quando se trata do conjunto dos significados, o *campo léxico* pode ser também denominado *campo semântico*. Cf. rede semântica.

código

Um *código* é um sistema de signos (ou sinais) que serve para transmitir mensagens e/ou informações. Ele é constituído por um conjunto de elementos e um conjunto de regras para combinar esses elementos. Ex.: o código de trânsito, uma língua qualquer.

cognato

Palavra que tem a mesma raiz de outra, ou outras, e com que constitui família de palavras. Ex.: *claro* e *clareza* são cognatos. Outros elementos dessa mesma família de cognatos: *claridade, claramente, clarear, clarificar, esclarecer*.

concordância

A *concordância* é um conjunto de citações ordenadas de uma dada palavra. A palavra pode aparecer à direita ou à esquerda da citação, ou contexto, e pode vir na forma invariável (subst. masc. sing.; verbo, infin.) ou nas suas formas flexionadas em ordem alfabética. As palavras que são postas em concordância seguem-se em listagem alfabética. As citações extraídas de obras são acompanhadas de códigos que permitem localizá-las no texto de origem. Existem concordâncias de autores, concordâncias de

obras (as *concordâncias* da bíblia, p.ex.). Há séculos zelosos lexicólogos vêm elaborando *concordâncias*. As bíblicas, por exemplo, são muito antigas. Hoje elas podem ser executadas automaticamente por um computador, a partir de um programa especialmente construído para este fim. Constituem uma fonte riquíssima de informações para a tarefa lexicográfica, poupando aos dicionaristas um gigantesco trabalho braçal na coleta de abonações para as entradas do seu dicionário. Em nossos dias as *concordâncias* textuais substituíram o arquivo do lexicógrafo nos grandes centros de Lexicografia.

conjunto aberto e conjunto fechado

Um *conjunto fechado*, ou *finito*, é um conjunto de elementos que podem ser circunscritos, enumerados. Em Linguística os fonemas de uma língua, os pronomes de um idioma constituem um *conjunto fechado*. Inversamente, um *conjunto aberto*, ou *infinito*, é aquele cujos elementos não podem ser perfeitamente enumerados e circunscritos. O léxico constitui um conjunto aberto. De fato, ele se caracteriza por sua permanente possibilidade de ampliação, à medida que uma cultura se expande e avança o conhecimento. O léxico só deixaria de ser um sistema aberto se a língua correspondente morresse, isto é, deixasse de ser falada por grupos sociais como ocorreu com o latim.

conotação — conotativo

Elemento da significação de uma palavra que não é comum a todos os falantes e pode variar segundo os contextos em que a palavra em questão está inserida. Um vocábulo como *luz* terá um valor semântico básico (denotação) e poderá ter inúmeros valores *conotativos* (abstrato, literário, emotivo, religioso etc.), dependendo do contexto em que ele aparecer e da pessoa que emite ou interpreta a mensagem.

corpus

Conjunto geralmente grande de dados lingüísticos que servem como matéria-prima para a pesquisa e a análise lingüística. Este termo começou a ser utilizado inicialmente na Estatística Léxica, onde os estudiosos levantaram grande volume de palavras em uma língua, para estudar o seu comportamento quantitativo, a sua distribuição na língua e no discurso. Hoje é usado em todos os domínios da Linguística com o significado acima definido.

definição ostensiva ou mostrativa

É um tipo de definição que utiliza a mostra do referente da palavra como recurso, em virtude da dificuldade de a linguagem verbal identificar claramente o *definiendum*. É muito freqüente nos dicionários. Ex.: *azul* = que é da cor do céu, do mar profundo, da água das piscinas. Esse tipo de definição torna-se muito apropriado quando o referente é um objeto, um ente, ou uma qualidade perceptível no mundo extralingüístico. O caso típico é o dos animais e plantas, em que os dicionários maiores e melhores geralmente mostram o referente da palavra através da sua imagem: um desenho, uma fotografia etc.

Cf. *definiendum*.

definiendum

A palavra-entrada do dicionário que deve ser definida pelo dicionarista.

definiens

A palavra (ou palavras) usada(s) para definir o *definiendum*.

denotatum — pl. denotata

Referente do signo lingüístico. No caso do signo *árvore*, a planta que se vê na natureza. Os *denotata* também podem ser abstratos sem correspondente no mundo extralingüístico: *constituição, sabedoria*. L.Zgusta (V. O dicionário unilingüe) usa muito estes termos (*denotatum, denotata*): a maioria dos lingüistas, porém, prefere usar referente.

designata

A coisa designada, a palavra designada, o que se refere, isto é, o referente.

dicionário bilíngüe

Dicionário que trata da correspondência das palavras entre duas línguas, procurando indicar a tradução de um termo de uma língua na outra língua considerada. Ex.: dicionário português-inglês, dicionário inglês-português.

dicionário unilingüe

Dicionário que contém as palavras de uma língua e as explica e/ou define através de paráfrases nessa mesma língua. Também chamado de *dicionário monolíngüe*. Contudo, a formação etimológica *unilingüe* é preferível por não ser heterogênea.

documentação

Enunciados, frases, parágrafos, onde está documentado o uso de uma palavra determinada numa obra literária ou científica, ou em qualquer outro tipo de escrito (jornalismo, p.ex.).

empréstimo

Palavra que foi tomada emprestada de uma outra língua por não existir termo correspondente no idioma que acolheu o empréstimo. Esse *empréstimo*, com freqüência, incorpora-se ao vocabulário vernáculo.

entrada

Cada uma das palavras explicadas por um dicionário. Em geral as entradas são alistadas em ordem alfabética e na forma canônica ou lema (Cf. essas palavras).

estatística léxica

Ciência que estuda o vocabulário do ponto de vista quantitativo. Os numerosos estudos estatísticos feitos sobre o léxico das mais variadas línguas evidenciaram que existem certas características estatísticas que aparecem em todas as línguas. A freqüência do uso de determinadas palavras (os instrumentos gramaticais, p.ex.) tem uma tal estabilidade que poderia ser considerada como um traço distintivo dessas palavras ao lado de outras características morfossintáticas. Outra característica que se repete em todas as línguas sobre as quais se fizeram estatísticas léxicas rigorosas é que as 1.000 palavras mais freqüentes de uma língua compõem mais de 80% de qualquer texto nessa língua. Esse índice pode chegar a quase 89%.

etiqueta

Rótulo qualificativo e explicativo que se atribui a cada entrada, ou a cada sentido, ou acepção de uma palavra. Essas *etiquetas* são geralmente abreviadas no contexto para não sobrecarregar o texto do verbete. Por exemplo: *abst.* (= significação abstrata);

met. (= significado metafórico); *col.* (= usado na linguagem coloquial); *reg.* (= regionalismo, só usado em determinada região); *arc.* (= arcaico, desusado, antigo) etc.

exemplário

Conjunto de exemplos (= frases, enunciados, contextos) utilizados para exemplificar o(s) significado(s) e o(s) uso(s) de uma palavra-entrada. Cf. abonação, documentação.

fonte

Autor ou obra que forneceu ao dicionarista um exemplo, ou uma citação com que ele abonou o significado e/ou o uso de uma palavra.

forma canônica

Forma de uma palavra considerada como forma-base para dar a ela entrada em um dicionário. As palavras que sofrem flexões morfossintáticas variadas e variáveis no discurso têm sido tradicionalmente representadas por uma das variantes do seu paradigma. Por exemplo: no caso do adjetivo, o masculino singular (*belo, bonito, amável, feliz*); no caso do verbo, a forma infinitiva (*amar, vender, partir*). Cf. lema.

glossário

Pequeno vocabulário, ou relação de palavras, em que se explica o significado das mesmas, para ajudar o leitor na compreensão do texto que lê. Modernamente são comuns os glossários de linguagem técnica como este.

índice de palavras

Também chamado *index-verborum*.

É uma lista das palavras que ocorreram em um determinado *corpus* (cf. essa palavra), ordenadas segundo dois critérios: 1) em ordem decrescente (ou crescente) da frequência com que ocorreram no *corpus*; 2) em ordem alfabética sendo indicada a frequência com que cada palavra ocorreu no *corpus*, à direita do vocábulo.

key-word-in-context

O mesmo que concordância (cf. essa palavra).

lema

Unidade léxica ideal que representa um paradigma de formas flexionadas. Essa unidade constitui a típica entrada de dicionário e representa todas as demais formas do paradigma. Por exemplo: a forma infinitiva *cantar* que aparece nos dicionários de português representa todas as outras variantes deste paradigma verbal: *canto, cantara, cantasse, cantando* etc. É, portanto, um *lema*. Cf. forma canônica.

lematização

O ato de reduzir uma forma que ocorreu num texto ao lema que passa a representá-la numa lista de palavras, ou num dicionário. Consideremos a oração: "Os dias correm céleres." A *lematização* dessa seqüência produziria os lemas: *o, dia, correr, célere*. Cf. forma canônica.

lexema

Unidade léxica abstrata que faz parte do léxico de uma língua. Ele se atualiza no discurso na forma de uma palavra flexionada com todas as marcas gramaticais exigidas pelo contexto. Cf. lexia, palavra.

lexia

Forma que um lexema assume no discurso. Ex.: “O dia está claro.” Temos aí quatro lexias. Em Lexicologia esse termo técnico se opõe a *lexema*; foi cunhado para evitar a imprecisão e ambigüidade de termos da língua comum, tais como: *palavra*, *vocábulo*.

lexicalização

O processo formal de categorizar, ao nível do léxico, uma forma, ou uma combinatoria de formas usadas no discurso. Ocorre lexicalização na evolução histórica de uma língua: 1) quando um *empréstimo* estrangeiro passa a fazer parte do léxico da língua (ex.: *coquetel*, *futebol*, *butique*); 2) quando uma criação vernácula nova é definitivamente aceita pela comunidade de falantes do idioma; 3) quando uma seqüência de formas passa a ser compreendida e usada como se fosse uma forma única, uma unidade léxica (ex.: *mercado negro*, *capa-de-chuva*).

lexicologia

Ciência que se ocupa do estudo do vocabulário de uma língua. Ela procura estudar o léxico enquanto sistema, e os seus elementos constitutivos nas suas peculiaridades. Opõe-se à Lexicografia que procura descrever o léxico, elaborando dicionários para consulta dos usuários da língua. A *Lexicologia* tem objetivos teóricos e especulativos. Tenta responder a questões como: o que vem a ser o léxico de uma língua quando confrontado com a gramática dessa mesma língua? como ele se estrutura? como ele se expande? como se dá o processo de categorização lexical? como se processa a aquisição e o enriquecimento do léxico de um indivíduo? como se faz o armazenamento do léxico nas memórias individuais? etc.

língua culta padrão

Modelo de linguagem considerado como padrão ideal numa comunidade lingüística, ou numa nação, se a língua for considerada relativamente às suas funções socioculturais e políticas. Coincide com a linguagem do grupo que possui maior competência lingüística e, por essa razão, goza de prestígio na sociedade: os intelectuais e os escritores.

memória do computador

Numa unidade central de processamento computacional existem vários equipamentos com finalidades distintas e com capacidade de operação diversa. Entre elas a *memória* principal que pode se subdividir em duas: 1) a *memória* que armazena os dados que serão analisados e processados; 2) a *memória* que armazena o programa (ou programas), isto é, a lista ordenada de instruções (algoritmos) que deverão ser seguidas na análise dos dados. Essas *memórias* podem ser ampliadas. Nos grandes centros de computação como, por exemplo, num sofisticado centro de pesquisa, ou num centro que opera milhões de dados, outras unidades de memória podem ser acopladas à memória principal a fim de ampliar a capacidade de armazenamento da máquina, tanto com relação a dados como com relação às instruções operatórias. Cf. algoritmo, analisador sintático e semântico.

metalingua

Um *metalingua* ou *metalinguagem* é um sistema de significação (uma linguagem), cujo objeto de descrição é um código e os seus signos. Um exemplo típico de *metalinguagem* é a definição e/ou descrição de uma palavra fornecida pelo dicionário.

Ex.: *crise* = / período difícil onde há dúvidas, incertezas e dificuldades antes de se poder chegar a uma solução que resolva a crise./ A *metalinguagem* dos dicionários usa os próprios signos da língua natural para explicar um signo do código lingüístico. Um outro tipo de *metalingua* (metalinguagem) se situa num nível mais abstrato. Opera com conceitos e signos lógico-matemáticos e se serve de uma teoria para analisar um objeto (uma língua natural, p.ex.), sem se reportar necessariamente a manifestações observáveis.

neologia

Segundo Louis Guilbert, grande especialista em Lexicologia, “a *neologia* lexical se define pela possibilidade de criação de novas unidades léxicas, em virtude das regras de produção existentes dentro do sistema léxico”.

neologismo

Unidade léxica (palavra) que é sentida como recente na língua pelos falantes do idioma. Existem *neologismos* em que o signo lingüístico todo é novo (forma e significado) — *jeans, fríser* — e outros em que só o significado é novo — *computador, átomo*. O estudo dos *neologismos* se relaciona diretamente com a história da língua; podemos estudar os *neologismos* do século XVI, os *neologismos* do século XIX, bem como os *neologismos* do século XX, ou das décadas 1950-80.

nomenclatura

A seqüência dos itens lexicais (as entradas do dicionário), ordenados geralmente em ordem alfabética, é o que os lexicógrafos chamam de nomenclatura.

obsoleto

Qualidade da palavra que não faz mais parte do vocabulário corrente, usual. Portanto, desusado, arcaico.

ocorrência

Unidade léxica que se atualiza no discurso produzido (oral ou escrito). Ex.: No verso de Fernando Pessoa “Ah, canta, canta sem razão” — temos duas ocorrências do lexema <cantar>

palavra

Unidade psico-sociológica fundamental da língua, essencial tanto no processo de comunicação, como no processo simbólico de apreensão do universo pelos sujeitos. É termo da língua comum, sendo pouco rigoroso para o uso técnico da Lexicologia e da Lexicografia. Charles Muller, em várias de suas obras, tentou atribuir aos termos *mot* (= palavra) e *vocable* (= vocábulo) valores técnicos na língua francesa para os seus trabalhos de Estatística Léxica (cf. essa palavra). A meu ver, essa tentativa se frustrou pelo mero fato de que essas palavras estão desgastadas pelo uso comum.

palavra instrumental

Palavra de valor e uso gramatical como preposições, conjunções, artigos etc. É um instrumento de articulação do discurso.

palavra lexical

Palavra de significação plena, não gramatical. Dentro do universo vocabular é

aquela que faz parte do conjunto de elementos exclusivamente lexicais, já que o outro conjunto é híbrido, pois contém elementos de natureza gramatical. Cf. palavra instrumental.

palavra plena

O mesmo que *palavra lexical*.

polissemia

Fenômeno lingüístico da acumulação de vários significados em um único lexema (palavra). Assim, um significante como *cavalo* pode significar: 1) animal quadrúpede e mamífero; 2) peça do jogo de xadrez; 3) tronco de uma planta onde se faz um enxerto. Toda língua possui um número enorme de palavras polissêmicas, pois esse fenômeno resulta do princípio da economia operante na língua. Em geral, quanto mais usada e mais antiga for uma palavra na língua, tanto maior costuma ser a pluralidade dos seus significados. É o caso de *ato*, *fazer* e *fundo*, p.ex.

recuperação de dados

Em Lingüística Computacional, Informática e Lexicografia, chama-se *recuperação de dados* o ato de retirar de um arquivo, ou de um banco de dados, determinados dados que se deseja num determinado momento.

Quando os dados estão estocados num disco que pode ser posto na memória do computador (cf. essa palavra), podemos recuperar informações que tinham sido anteriormente armazenadas no disco, através de um programa que dá as instruções adequadas ao computador.

recuperação de informação

O mesmo que recuperação de dados.

rede semântica

Uma rede semântica é um conjunto articulado e estruturado de campos *léxicos* (cf. essa palavra) que se comunicam entre si. Os campos léxicos vão-se integrando uns nos outros, através de nós sêmicos comuns (unidades de significação), a fim de compor a grande cadeia de uma rede semântica.

referência cruzada

No texto de um verbete (entrada de dicionário) é freqüente o dicionarista remeter a outra palavra. Essa prática se explica da seguinte forma: além de economizar espaço no dicionário, evita-se repetir informações que já foram dadas em outro verbete e por isso se remete a ele. Por outro lado, para que o consulente compreenda bem o significado e o uso de uma palavra, é preciso contrapô-la a outras palavras de significação próxima, ou oposta. Ninguém aprende ou apreende uma unidade de léxico isoladamente, uma vez que o léxico é formado de campos léxicos e grandes redes semânticas integradas. Muitas vezes, é só através da compreensão de um campo léxico, ou de uma rede, que o usuário poderá tomar consciência exata do valor de uma palavra.

referente

Em Lexicologia e Semântica chama-se de *referente* o objeto ou conceito que é expresso através de um signo lingüístico (uma palavra ou uma locução); a coisa, a noção à qual é feita referência. Ex.: o *referente* de *cavalo* (1.º significado, cf. polissemia) é o

animal quadrúpede; o *referente* de *saudade* é o sentimento que se experimenta e que essa palavra refere.

registro

Um registro lingüístico é uma variante da língua (com gramática e vocabulário próprios), utilizado em função de determinados fatores humanos e sociais. São fatores determinantes para a seleção do registro a ser usado pelo falante: 1) o contexto e a situação; 2) os interlocutores em presença; 3) o tema do discurso. Os registros derivam da natureza social da linguagem e das atitudes sociais dos indivíduos em relação à língua.

repertório léxico (vocabular)

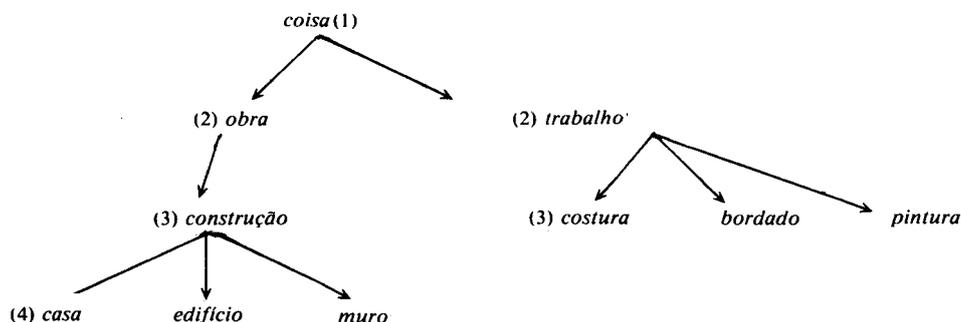
Conjunto das palavras de uma língua, ou universo léxico. Pode-se falar também no *repertório léxico* (vocabular) de um indivíduo, de uma obra, de um grupo social, dando a esse termo um significado mais restrito.

semasiologia

A *semasiologia* é uma área da Semântica que estuda os significados e a sua estruturação interna, a partir dos signos lingüísticos (das palavras). Divide-se assim o signo lingüístico nas suas duas faces — significante e significado. O método semasiológico considera os significantes para indagar sobre os significados, ou investigar o fenômeno da significação. O contrário da *semasiologia* é a *onomasiologia*, que parte da significação em busca da designação lingüística dos conceitos ou objetos considerados.

superordenado

Os lexicólogos admitem que o léxico de uma língua está estruturado não só em campos léxicos (cf. essa palavra) e redes semânticas (cf. essa palavra), mas também de maneira hierárquica. De acordo com essa última característica, o vocabulário se ordenaria em cadeias hierarquizadas, onde alguns termos estão situados teoricamente acima de outros, isto é, são superordenados em relação aos que estão mais abaixo na cadeia em árvore. Ex.:



(1) é superordenado em relação a (2), (3) e (4)

(2) é superordenado em relação a (3) e (4)

(3) é superordenado em relação a (4)

tesouro ou thesaurus

Dicionário completo que deveria conter todas as palavras de uma língua, descrevendo o universo vocabular desse idioma na sua totalidade.

Não só a sua nomenclatura deverá ser exaustiva, como a redação dos verbetes, incluindo-se neles todas as denotações e conotações possíveis da palavra-entrada. Ainda mais: deverá conter uma exemplificação e documentação completa dos valores e usos de cada palavra; para tanto, a equipe responsável pela elaboração de um *thesaurus* deverá compilar todos, ou quase todos, os textos escritos da língua para evitar omissões. Uma tarefa dessa é evidentemente gigantesca. Só é imaginável nos tempos contemporâneos quando grandes equipes de especialistas se reuniram em centros de pesquisa especialmente criados para este fim, como é o caso do *Trésor de la langue française*, que se elabora há mais de vinte anos em Nancy (França) com o concurso de computadores e de grandes verbas do governo francês. No caso de uma língua morta como o latim, todo dicionário deveria ser do tipo *thesaurus*, já que o universo léxico dessa língua se fechou, pois não existem mais falantes desse idioma e, portanto, não se criam mais neologismos (cf. essa palavra).

verbeta

O texto de uma palavra-entrada de um dicionário, inclusive ela própria. Os dicionários são formados de seqüências de verbetes.

vocábulo

Sinônimo de *palavra* na língua comum; portanto, é ambíguo e não tem os requisitos de rigor técnico necessários à linguagem científica, a saber: um significado unívoco. Para completar as informações sobre *vocábulo* consulte o verbeta *palavra*.

unidade léxica

Unidade vocabular tanto com respeito à significação como com respeito à forma gramatical, que tem um uso característico no discurso. Sinônimo de *lexema* (cf. essa palavra), ou de *lexia* (cf. essa palavra) conforme o contexto.

usuário

1) O consulente de um dicionário; o indivíduo que usa o dicionário para informar-se sobre o significado e o uso de uma palavra. 2) O falante da língua, aquele que usa a língua para comunicar-se.