

PROPOSTA PARA A CONSTRUÇÃO DA MICROESTRUTURA DE UM VERBETE TERMINOLÓGICO PARA TRADUTORES

*Guilherme Fromm**

RESUMO: o presente artigo apresenta um possível modelo para a construção de microestruturas de obras terminográficas para tradutores, tendo como sua única fonte corpora textuais em diferentes áreas e previamente construídos. Também se discute a construção da parte mais difícil de uma microestrutura – a definição – e propõe padronização para os paradigmas da microestrutura.

UNITERMOS: lexicografia; terminografia; tradução; linguística de *corpus*.

ABSTRACT: the present article shows a possible model of microstructure construction in terminographical works for translators having text corpora, in different areas and previously built, as its only source. It also discusses the construction of the most difficult part of a microstructure, the definition, and proposes standardization for the microstructure paradigms.

KEYWORDS: lexicography; terminography; translation; corpus linguistics.

1 Introdução

Uma das primeiras preocupações de um lexicógrafo ou terminógrafo é o estudo que fará para determinar a construção da

* Doutor em Letras pela FFLCH/USP, professor Adjunto do ILEEL/UFU.

microestrutura de sua obra tendo em vista o público-alvo da mesma. O objetivo deste artigo é a apresentação de um modelo de microestrutura para um dicionário terminológico, mono ou bilingue, voltado aos tradutores.

Partindo dos conceitos de Haensch (1982:470-501), podemos dividir a inserção de dados na microestrutura em diferentes seções as quais denomino “Paradigmas Terminográficos”. Eles representam, na microestrutura de um termo, os campos que a compõem. O autor já fazia uma divisão dos componentes básicos da microestrutura de um verbete lexicográfico e os denominava Paradigmas dos tipos: Informacional (informações ortográficas, fonéticas, etimológicas, cronológicas, gramaticais etc.); Definicional (definição); e Pragmático (exemplos e abonações). Segundo Barbosa (1994), a organização dos paradigmas que compõem a microestrutura está relacionada com a obra lexicográfica (ou terminográfica, no caso) em que aparecem. Não há uma fórmula pronta para a composição de uma obra lexicográfica ou terminográfica; o que deve ser observada, desde a fase de elaboração, é a coerência interna da mesma. Ainda, para a autora, (1989:571), novos paradigmas podem ser acrescentados para o enriquecimento da obra (não há, portanto, um número “correto” de paradigmas), desde que o principal deles, o Definicional, não seja deixado de lado.

As fontes dos dados que preencherão a microestrutura podem advir de vários meios: outros dicionários, enciclopédias, obras especializadas etc. A proposta aqui é trabalhar com a Linguística de *Corpus*¹ e utilizar dados selecionados de *corpora* construídos especificamente para cada obra terminográfica a ser produzida². Para tanto, criou-se um banco de dados eletrônico, que pode ser preenchido via Internet (disponível em: <http://www.guifromm.trd.br/admin>; acesso restrito aos pesquisadores cadastrados).

¹ Um *corpus*, segundo Tagnin (2004), é “[...] uma coletânea de textos em formato eletrônico, compilada segundo critérios específicos, considerada representativa de uma língua (ou da parte que se pretende estudar), destinada à pesquisa”.

² A descrição de como esses *corpora* foram construídos pode ser verificada em Fromm (2008).

Esse banco aceita a inclusão de contextos (explicativos ou definitórios, discutidos por Aubert, 1996) previamente selecionados nos *corpora* (pois o banco pode ser trabalhado de modo mono ou bilingue) e, a partir desses contextos, buscam-se os dados para preencher as estruturas previamente selecionadas. Os campos disponibilizados para o preenchimento dos paradigmas também são fruto de duas pesquisas previamente realizadas: uma cotejou a bibliografia especializada, a outra é resultado de uma pesquisa com mais de 170 tradutores para conhecer suas preferências no uso de dicionários.

São apresentados, a seguir, os Paradigmas que compõem a microestrutura proposta para este modelo.

2 O Paradigma Pragmático (PP)

A tela do banco de dados solicita, para cada termo, tantos contextos quanto o pesquisador achar necessários para a alimentação dos demais campos e a construção da definição; esses contextos são extraídos do *corpus*. Cada contexto exige o preenchimento de quatro campos: exemplo (o contexto em si), fonte, data de coleta e conceito (figura 1-A)³.

Fig. 1-A: Tela de cadastro de contextos, parte superior – termo *memória*

³ Todos os exemplos aqui apresentados foram tirados de um corpus bilingue de informática, por mim coletado, contendo cerca de um milhão de palavras cada língua.

A partir de cada contexto, o pesquisador deve, também, extrair um primeiro conceito formado pelos traços distintivos explicitados pelo exemplo (e que serão retomados na construção do Paradigma Definicional). Veja o exemplo, na figura 1-B, com o termo “memória” (o termo é bastante polissemêmico, não só nessa área como em outras; a especificação do mesmo na área *hardware*, nesse caso, demandou uma grande quantidade de contextos retirados do *corpus*):

Contextos Cadastrados			
Exemplo	Conceito	Fonte	Ações
Quando o micro é ligado o processador carrega o sistema operacional, executa programas e manipula os dados conforme solicitado pelo usuário. Sabendo que, o processador não tem uma área interna de armazenamento muito grande de dados, essas informações são armazenadas em uma área chamada memória.	área para armazenamento de dados	CDH 08/06/2007	editar - excluir
Um chip de memória é um exército de clones, formado por um brutal número de células idênticas, organizadas na forma de linhas e colunas, de uma forma similar a uma planilha eletrônica. O chip de memória em si serve apenas para armazenar dados, não realiza nenhum tipo de processamento.	chip formado por células idênticas, na forma de linhas e colunas; apenas armazena dados, não realiza nenhum tipo de processamento	GHD 08/06/2007	editar - excluir
A maneira mais rápida de analisar o desempenho da memória é compararmos a sua taxa máxima de transferência, que é calculada com a fórmula $\text{clock} \times \text{bits por clock} / 8$. Os chips voltados ao mercado de alto desempenho acessam a memória a 256 bits, contra 128 bits ou mesmo 64 bits dos chips voltados aos mercados intermediário e iniciante.	desempenho medido pela taxa de transferência (256, 128 ou 64 bits).	CDH 01/02/2005	editar - excluir
O ambiente de edição é o que mais exige recursos do computador. Isso pode ser notado com o uso de um micro com chip Celeron 2,8 GHz e 256 Mbytes de memória RAM.	medida em unidades de Mbytes	FSP 01/02/2005	editar - excluir
Quando a memória RAM está ocupada, o Windows cria a chamada memória virtual, ou seja, usa o disco rígido, que é muito mais	pode ser virtual, criada pelo windows	FSP 01/02/2005	editar - excluir

Fig. 1-B: Tela de cadastro de conceitos, parte inferior. Aqui são apresentados os contextos e respectivos conceitos já cadastrados para o termo *memória*

Esses exemplos são os mesmos que serão disponibilizados ao consulente na página de consulta proposta, denominada VoTec (disponível em: <http://jr.icmc.sc.usp.br/~comet/dic/>; site funcional, mas a quantidade de verbetes disponíveis ainda é reduzida). No modo padrão, somente um exemplo é mostrado; mas o usuário pode solicitar a visualização de todos ao manipular as opções da página.

A próxima tela, após a entrada dos contextos, disponibiliza uma segunda página de inserção de dados sobre o termo (as figuras seguintes apresentam as várias abas que constituem essa página) e servirá para a elaboração de todos os demais paradigmas para a construção da microestrutura do termo.

3 A construção do Paradigma Definicional (PD)

A elaboração da definição (“tipo específico de texto bidimensionado: com suas dimensões semântica e pragmática” [Finatto, 1994]), como em qualquer obra lexicográfica ou terminográfica, representou um desafio para a construção do banco. Em primeiro lugar, há de se apresentar uma distinção entre a definição lexicográfica e terminográfica tradicional:

A diferença entre a definição terminológica e a definição lexicográfica remete, portanto, a perspectivas distintas de abordagem, muito embora elas não sejam isentas: a lexicografia parte do signo para chegar à determinação do conceito (procedimento semasiológico); a terminografia parte da noção (ou conceito) e pesquisa os termos que lhe correspondem (procedimento onomasiológico). (Lara, 2000:94)

Vários outros autores trabalham com a questão da terminografia moderna e suas implicações. Finatto (1998) faz um bom resumo do que é a criação da definição terminológica:

No âmbito do dicionário terminológico, tal como verificamos, via de regra, na definição contida na microestrutura, definir pode ser o mesmo que discriminar características, explicar, situar ou arrolar diferentes modalidades ou variedades do definiendum, corporificado pelo termo, o que se dá pela expressão lingüística de um conceito de estatuto técnico ou científico. Entretanto, a formulação da definição estará igualmente mediada e condicionada por vários elementos tais como a natureza da obra, pelo tipo de discurso em que originalmente se inscrevem o conceito e o termo, pelo tipo de público-alvo e perfil do produto terminográfico em que se insere, entre outros fatores. (Finatto, 1998)

A premissa básica para a construção do banco de dados foi que todos os campos a serem trabalhados deveriam ser preenchidos a partir de, e unicamente a partir de, contextos retirados de *corpora* reais de língua. Para tanto, houve uma grande preocupação com o balanceamento da quantidade de palavras dentro de cada subárea do

corpus para que o processo de construção das definições fosse o mais acurado possível, já que “(...) a microestrutura de dicionário, em sua porção definicional, é um tipo de DT⁴ em que encontramos, via de regra, a reescritura de um outro texto, o texto de referência a partir do qual é, geralmente, moldada, extraída ou adaptada...” (Finatto, 1998).

Os contextos definitórios e explicativos foram buscados já durante a elaboração desses *corpora* (o que não significa, necessariamente, que todos os textos os possuam). Os textos reunidos representam a seleção de dois tipos de linguagem: de especialistas para especialistas (textos acadêmicos e manuais técnicos de alguns produtos) e de especialistas para leigos (textos jornalísticos explicativos, com características de enciclopédia, e outros manuais técnicos).

A definição, partindo de uma abordagem linguística, cognitiva e comunicativa⁵ pode, então, ser composta pela reescrita desses dois tipos de textos. E isso para um público-alvo que terá de lidar com as mesmas questões nas traduções do seu dia a dia: textos técnicos de especialista para especialista, textos jornalísticos de especialista para leigos etc.

Alves (1996:125), ao discutir as propostas de Sager (1990), um entre os vários autores que tratam do assunto (Almeida; Souza; Pino, 2007, fazem um ótimo apanhado daqueles que trabalham no tema), nos apresenta os diversos tipos de construção da definição em si:

- a. Analítica (ou por compreensão): termos genéricos e características específicas;
- b. sinonímicas;
- c. paráfrase;
- d. síntese: descrevem e identificam relações;
- e. implicação: uso do termo em um contexto explicativo;
- f. denotação (extensão): relacionam exemplos;
- g. demonstração: definem por meio de fotografias, ilustrações, etc.

⁴ DT, no texto da autora, remete à Definição Terminológica.

⁵ Segundo Finatto (1998), há dois tipos de abordagem para os fenômenos terminológicos: uma linha afinada com as ideias de Sager, com enfoque linguístico, e uma linha tradicional, com enfoque em estudos ontológicos e na Teoria do Conceito.

A definição é um “enunciado que descreve um conceito e permite sua diferenciação de outros conceitos dentro de um sistema de conceitos”⁶; já conceito é “uma unidade de pensamento constituída através de abstração sobre a base de propriedades comuns a um conjunto de objetos”⁷ (ISO 1087, 1990:1; minhas traduções). A norma ISO 1087 (1990:4) propõe dois tipos básicos de definição:

4.1 definição por compreensão: definição baseada na compreensão⁸ de um conceito (Nota: para este propósito é necessário especificar o conceito genérico mais próximo que já tenha sido definido ou que possa ser assumido como de conhecimento geral e adicionar as características de restrição que delimitam o conceito a ser definido).

4.2. definição por extensão: definição baseada na enumeração exaustiva de objetos que se referem ao conceito, ou de conceitos específicos no próximo nível de abstração.⁹

Baseada nessa norma, Alves (1996) sustenta que

[a] definição por compreensão (intensional) é, certamente, a mais adequada aos trabalhos terminológicos, uma vez que ela situa o conceito no âmbito de uma classe e especifica o que o distingue de outros conceitos situados na mesma classe. (Alves, 1996:126)

⁶ *Statement which describes a concept and permits its differentiation from other concepts within a system of concepts.*

⁷ *A unit of thought constituted through abstraction on the basis of properties common to a set of objects.*

⁸ [Conjunto de características que constitui um conceito.] *Set of characteristics which constitutes a concept* (ISO 1087, 1990:2).

⁹ 4.1 intensional definition: Definition based on the intension of a concept (Note: for this purpose it is necessary to state the closest generic concept that has already been defined or can be assumed to be generally known, and to add the restricting characteristics that delimit the concept to be defined); 4.2 extensional definition: Definition based on the exhaustive enumeration of the objects referred to by the concept, or of the specific concepts at the next level of abstraction.

Os termos genéricos que nela aparecem não pertencem a um domínio em particular e introduzem o conceito do termo que está sendo descrito. Como exemplo desses termos, temos: elemento, estrutura, fase, fórmula, método, parte, procedimento, processo, sistema, símbolo, subárea, técnica etc.

Gianni¹⁰ (1993, *apud* Finatto, 1994) acrescenta à norma 1087 a “definição contextual”, que reproduz o contexto em que o termo ocorre, e a definição enciclopédica, orientada à descrição exhaustiva da coisa nomeada. A construção da definição proposta para os verbetes não inclui esses elementos, mas eles reaparecem na composição da microestrutura (através dos Paradigmas Pragmático e Enciclopédico), ampliando os horizontes semânticos do termo em questão.

Na figura 2 abaixo, para o termo *processor*, por exemplo, podemos verificar como se dá a construção da definição no banco. Essa página representa a segunda página de inserção de dados. Nela são visualizados, num primeiro momento, na parte superior, os exemplos extraídos pelos pesquisadores na primeira tabela e seus respectivos conceitos iniciais; a partir dos mesmos, os traços distintivos serão agrupados em linhas (parte inferior). O terminógrafo deverá agrupar traços que envolvam as mesmas ideias, nas diferentes linhas, nas mesmas colunas (como *performs* e *execute*, no exemplo abaixo).

The screenshot shows a web browser window with the URL <http://prcos.sc.usp.br>. The page title is "Termo: processor". There are navigation buttons: "Passo Anterior", "Salvar", and "Sair sem salvar". Below these are tabs for "Contextos" and "Conceito Final / Definição".

The main content area is divided into three columns: "Exemplo", "Conceito", and "Fonte".

Exemplo	Conceito	Fonte
1 The "brain" of any computer is the central processing unit (CPU). This component does all the calculations and performs 90 percent of all the functions of a computer. There are many different types of processors for computers—so many, in fact, that you will learn about them later in this chapter in the section "Identifying Purposes and Characteristic of Processors."	does all the calculations and performs 90% of all the functions of a computer	Wiley 08/06/2007

Below this is a section for "Traços Distintivos" with a sub-header "Nove Coluna" and a note "— Não use aspas duplas". It contains a table with columns A through I and rows 1 through 4.

	A	B	C	D	E	F	G	H	I
1	calculates	performs	functions						
2				flat squar...	circuits e...				
3						speed mea			
4		execute					CPU	control	programs

Fig. 2: Contextos, conceitos e seleção de traços distintivos

¹⁰ GIANNI, E. (1993) O Paradigma definicional lexicográfico e terminográfico. *Cadernos do IL*. Porto Alegre: UFRGS, n°. pp. 45-56.

Desses traços distintivos, oriundos dos contextos do *corpus*, serão tirados o conceito final e a definição do termo. O conceito final é montado tendo em vista os vários traços agregados nas diferentes linhas e colunas (note-se que esses traços, ao comporem a definição, podem não estar na mesma ordem em que aparecem nas linhas e colunas). O trabalho do terminógrafo, aqui, pode representar variações na construção do conceito final (diferentes pesquisadores percebem diferentes traços ou os agrupam de formas variadas).

O ponto de partida adotado para a construção do conceito final é o uso do paradigma GPDE (gênero próximo, diferença específica) adotado pelo GETerm (Almeida; Souza; Pino, 2007), ou seja, a definição começa pelo termo superordenado (hiperônimo) em relação àquele que está sendo descrito e, em seguida, são apresentados os semas específicos do termo em questão.

O termo superordenado (normalmente fornecido pelos exemplos do *corpus*) é imediatamente superior ao termo em questão e não representa, necessariamente, uma subárea da árvore de campo. Quando os excertos do *corpus* não forneceram o termo superordenado, foram usados dois termos genéricos dentro da área de hardware aqui descrita: *máquina* ou *equipamento* ou ainda descritores mais genéricos, como *tipo de* (produzindo enunciados como *verbete x: é um tipo de máquina que...*). Como a maioria dos termos levantados a partir do *corpus* são substantivos, estipulou-se que o primeiro descritor também deveria ser um substantivo. Continuando com o exemplo da figura 4, em português, tem-se para o termo *processador* a seguinte definição:

tipo de chip, parte principal do computador, também conhecido como CPU (Central Processing Unit); executa instruções (processos) e dados (através da UC - unidade de controle) e operações lógicas e aritméticas (através da UAL - unidade aritmética e lógica).

Outra decisão tomada pelo GETerm, e aqui também adotada, foi a quantidade de semas que compõem a diferença específica. Como a elaboração do conceito final representa uma pluralidade de concepções sobre o termo, retiradas dos vários excertos do *corpus*, decidiu-se pela agregação do maior número possível de semas na sua definição, mesmo que essa tivesse um caráter parcialmente enciclopédico. Em virtude disso, ao contrário do que

propõem alguns pesquisadores, como Dubuc¹¹ (1999, *apud* Almeida; Souza; Pino, 2007), sobre a construção da definição, a mesma não obedeceu, necessariamente, ao preceito da concisão. Campos (1994:46) comenta que quanto mais curta e concisa é uma definição, mais difícil ela se torna. Como no exemplo acima apresentado, quando da construção de um conceito com mais de uma sentença, preferiu-se, no intuito de tornar o entendimento mais claro, o uso de ponto-e-vírgula em vez de longas sentenças separadas apenas por vírgulas.

Tem razão Fraústo ao afirmar que “(...) a construção de uma terminologia, informatizada ou não, implica uma forte componente criativa” (Fraústo, 1990:1). “Engessar” a construção do conceito final, elaborado a partir de um *corpus* que sempre pode ser atualizado, limita as possibilidades de reconstrução desse conceito. Além disso, qual a garantia de que normas internacionais referentes à terminografia hoje vigentes serão adequadas para essas novas composições de *corpora*? A preocupação primeira do terminógrafo, então, deve ser com o seu público-alvo. Por essa razão, tendo em vista os tradutores e a dinamicidade da língua, preferiu-se trabalhar com esse componente criativo: não há razão para “engessar” a construção da definição.

É importante notar, também, a diferenciação entre conceito final e definição, no banco de dados, como mostrado na figura 3 abaixo. Cabré (1993) propõe que a definição deva obedecer aos critérios previamente estabelecidos quando do planejamento da microestrutura da obra. A definição, no presente caso, pode refletir esses critérios e modificar (ou não) os dados fornecidos pelo conceito final, adequando o texto final da definição às características da microestrutura da obra em desenvolvimento. No caso do campo de computação, neste trabalho, decidiu-se pela duplicação, em todos os casos, do conceito final na definição. Levando em conta o componente criativo, não se constatou a necessidade de construção de uma definição totalmente igual à construção de outra, ou seja, adotar a padronização e de forma generalizada, isto é, para todos os verbetes.

¹¹ DUBUC, R. (1999) *Manual práctico de terminología*. Chile: RiL Editores.

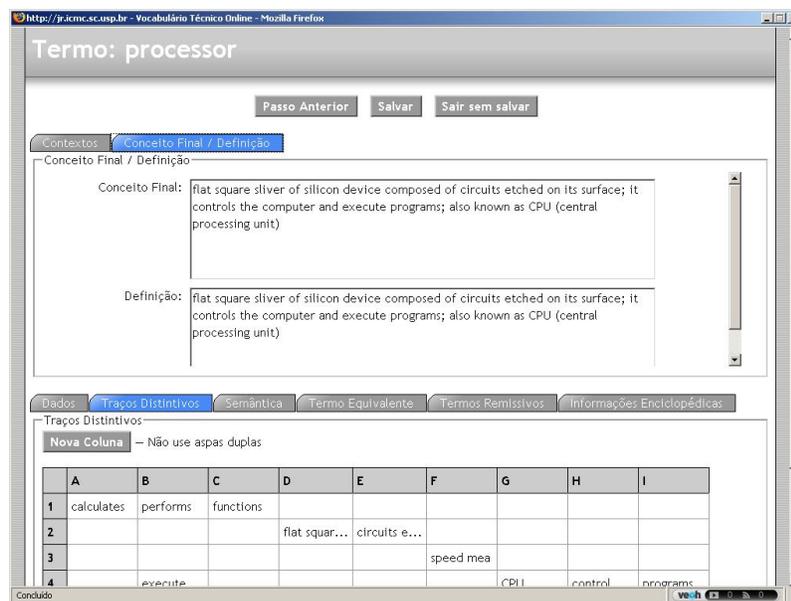


Fig. 3: Os traços distintivos (parte inferior) e a criação do conceito final e definição (parte superior)

Barbosa (1994) já alerta que um conjunto noêmico (conceito) de uma língua e seu respectivo conjunto léxico-semântico nem sempre encontram os equivalentes perfeitos em outra língua, demandando, por parte do lexicógrafo/terminógrafo, uma nota explicativa de caráter enciclopédico. Uma característica importante na terminografia, e que vai ao encontro dessa observação de Barbosa, pode ser encontrada nas definições levantadas nas duas línguas: elas não são, necessariamente, uma a tradução da outra; representam, na verdade, o processo de descrição, em cada língua e em cada momento, de um determinado conceito. Têm razão González-Jover e Sierra (2007) quando mencionam que

[...] o dicionário não só há de lhe oferecer aquela informação que se considera necessária e suficiente sobre os conceitos, mas também deverá incluir informação “enciclopédica”, relativa ao conhecimento extralinguístico. Nesse sentido, o dicionário especializado bilíngue destinado ao tradutor tem que conter informações que podem

superar os limites do estritamente necessário e suficiente: há de fazer as vezes de um dicionário monolíngue, no qual apareçam uma definição e contextos de usos real do termo validado que assegurem a confiança da informação e, assim mesmo, há de conter outras informações como sinônimos, variações significativas determinadas pelo contexto, categoria gramática, relações conceituais, etc., que sirvam de ajuda ao tradutor no processo de tomada de decisões que é a tradução. (González-Jover e Sierra, 2007:4, tradução minha)¹²

Em outras palavras, o levantamento dos campos da microestrutura, no banco de dados, foi pensado como monolíngue: o que os une, como dicionário bilíngue, são os Paradigmas de Forma Equivalente e a visualização na página de consulta (VoTec). Com essa abordagem, o banco de dados pode funcionar, também, para a construção de vocabulários monolíngues. A diferenciação no modo de consulta se dará a partir da criação de diferentes páginas (*sites*), com diferentes características de busca, que acessam o banco.

Vejamos exemplos de definições bilíngues de alguns termos que estão disponíveis para consulta na página do VoTec:

1. computador (*computer*): máquina capaz de realizar tarefas de tratamento de informações em alta velocidade,

¹² [...] el diccionario no sólo ha de ofrecerle aquella información que se considera necesaria y suficiente sobre los conceptos, sino que también deberá incluir información de la denominada «enciclopédica», relativa al conocimiento extralingüístico. En este sentido, el diccionario especializado bilingüe destinado al traductor ha de contener informaciones que pueden superar los límites de lo estrictamente necesario y suficiente: ha de hacer las veces de un diccionario monolingüe en el que aparezcan una definición y contextos de uso real del término validados que aseguren la fiabilidad de la información y, asimismo, ha de contener otras informaciones como sinónimos, variaciones significativas determinadas por el contexto, categoría gramatical, relaciones conceptuales, etc., que sirvan de ayuda al traductor en el proceso de toma de decisiones que es la traducción.

com precisão, formada por circuitos e dispositivos./ *device with a microprocessor that receives inputs, processes them and displays results.*

2. impressora (printer): periférico composto por placa lógica, motor, engrenagens, correias, eixo, carro de impressão e cabeça de impressão; imprime fotos, desenhos, textos, tabelas, dados, relatórios etc.; podem ser coloridas ou branco e preto; os três tipos principais são: matricial, jato de tinta e laser./ *peripheral device that prints two-dimension graphics, texts, color charts and photographs in hard-copy paper; the two main types are laser and inkjet.*
3. memória (memory): chip (formado por células idênticas, na forma de linhas e colunas) ou área para armazenamento de dados; unidade exibida em Mbytes, cujo desempenho é medido pela taxa de transferência; pode ser real ou virtual, associada ao Windows/ *place where the computer stores current programs and data (information) in use; it's connected to the processor by its address and data buses.*

Os exemplos mostram que as definições, em ambas as línguas, não são idênticas e oferecem semas complementares para o melhor entendimento do conceito. Isso se dá em virtude de uma das particularidades do campo da computação: a língua em que uma nova tecnologia ou processo foram criados em primeiro lugar. Os textos criados para a confecção de manuais ou os artigos jornalísticos representam, em determinada língua e em determinado momento, o estágio de desenvolvimento da técnica. Textos em outras línguas, sobre o mesmo assunto, podem ter sido produzidos na mesma época ou não. Esse é um dos fatores (o outro seriam as diferenças, várias, entre as fontes) que gera as diferenças nas definições: enquanto os textos de um termo na língua X, na qual o produto/tecnologia/processo foi criado há alguns anos, apresentam poucos semas para a construção da definição, outros textos, na língua Y, que descrevem esse produto/tecnologia/processo hoje, apresentam uma quantidade maior de semas.

Há claramente uma tendência a vulgarizar (passagem da terminologia para a língua comum; Barbosa, 1998) os termos com

o passar do tempo; e eles, quando da sua apresentação nos textos, passam de um tipo de contexto para outros: de contextos definitórios, tornam-se explicativos e depois associativos, tal como o termo *computador*. Não há a necessidade, por exemplo, de se explicar o que é um computador em países nos quais a tecnologia já existe há décadas; necessidade haverá em países onde a tecnologia ainda é incipiente. No caso do *corpus* de computação, significa que termos mais antigos, em inglês, podem apresentar, nos contextos, menos semas para a construção da definição do que os equivalentes, em português.

4 Paradigma Informacional (PI)

O próximo passo, ainda na mesma janela (parte inferior), com a visualização dos contextos em destaque, é a extração de informações morfossintáticas (figura 4): classe gramatical, número, gênero e possíveis siglas ou acrônimos¹³, entrada por extenso (caso o termo seja uma sigla ou acrônimo), variações morfossintáticas, o número da acepção¹⁴. São inseridos também os campos que localizam o termo em relação ao *corpus*: a frequência do termo na listagem de todas as palavras e o número total de vezes em que ele aparece.

¹³ “Muitos dos termos sintagmáticos, por razões de economia discursiva, são reduzidos sob a forma de siglas, termos formados pelas letras iniciais de cada elemento do sintagma, ou de acrônimos, termos que se constituem com a formação de sílabas extraídas do sintagma, geralmente as iniciais” (Alves, 2007).

¹⁴ Embora obras terminológicas tendam a apresentar definições monossêmicas, preferimos inserir esse campo. Algumas áreas, que já apresentam uma grande diversidade de terminologia, como a informática, podem vir a apresentar alguns termos polissêmicos em suas diferentes subáreas. Embora a definição do termo seja única em cada uma dessas subáreas (ou seja, são monossêmicas), numa consulta genérica à área como um todo, mais de uma definição poderia ser apresentada.

Dados

Ontologia: Computação > Hardware

Categoria Gramatical: Substantivo Número: Singular

Gênero: Masculino Sigla/Acrônimo: CPU, UCP

Entrada por Extenso:

Var. Morfossintáticas:

Aceção Nº: 1

Cópus

Posição na Ordem de Freqüência: 67 Nº de Ocorrências do termo: 1212

18/07/2007 15:59 © 2007 FFLCH - ICMC Jr.

Fig. 4: Informações morfossintáticas do termo “processador”

5 Paradigma Semântico (PS)

O Paradigma Semântico expõe as relações de hiperonímia, hiponímia, co-hiponímia, antonímia e sinonímia do termo em questão com outros termos disponíveis nos contextos extraídos, caracterizando possíveis remissivas. Além disso, o pesquisador poderá cruzar referências com obras já publicadas, verificando se o termo é dicionarizado, se há definições coincidentes, a fonte da definição e a definição dicionarizada em si. A tela abaixo (figura 5) apresenta as possibilidades acima descritas para o termo “processador”:

Semântica

Termo Dicionarizado? Definições Coincidentes? Parcial

Fonte: Dicionário eletrônico Houaiss da língua portuguesa

Definição Dicionarizada: circuito integrado capaz de efetuar o processamento de dados, obedecendo a um conjunto predeterminado de instruções

Hiperônimo de: UAL (unidade aritmética e lógica); UC (unidade de controle)

Hipônimo de: computador

Co-hipônimo de: chip

Sinônimo(s): CPU, UCP Antônimo(s):

Notas:

18/07/2007 15:59 © 2007 FFLCH - ICMC Jr.

Fig. 5: Informações semânticas e referência do mesmo termo/lexema em outras obras, termo *processador*

6 Paradigma de Forma Equivalente (PFE)

Nesta proposta bilingue, cada termo na língua de partida deve ter o seu equivalente na língua de chegada. É o preenchimento desse campo (figura 6), conforme descrito anteriormente, que possibilitará a visualização dos vocabulários como bilíngues (o resultado pode ser analisado no próprio site de consulta, o VoTec). Os equivalentes, já detectados na fase de análise dos *corpora*, precisam ser, agora, apenas associados.

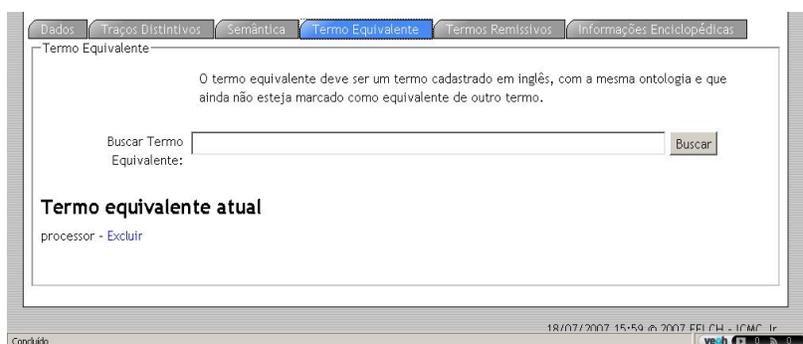


Fig. 6: Termo equivalente a “processador”, em inglês (em destaque)

7 Remissivas

Termos remissivos (remissivas são relações, traçadas dentro da obra lexicográfica/terminológica, entre os termos, mantendo a coerência semântica dessa [Fromm, 2002]), levantados a partir dos campos semânticos acima citados (sinônimos, antônimos, hipônimos, co-hipônimos e hiperônimos), só poderão ser incluídos num verbete, se esses mesmos verbetes já fizerem parte do banco como outros verbetes. Embora sejam discriminados, também, na composição da microestrutura, eles só funcionarão como *hyperlink* nessa mesma microestrutura se forem inseridos (figura 7) como elementos de pesquisa remissiva.

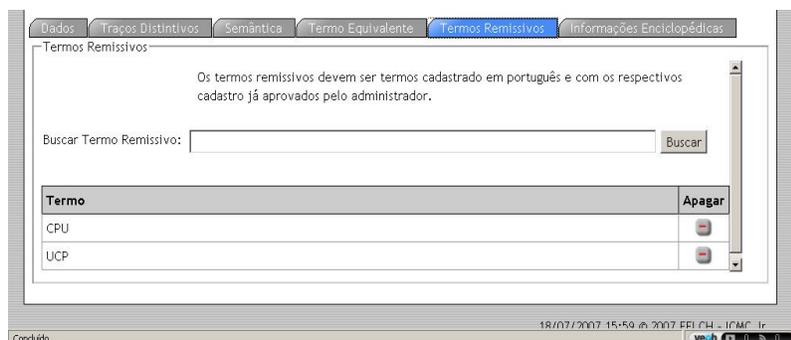


Fig. 7: Termos remissivos para *processador*

8 Paradigma Enciclopédico (PE)

Há a possibilidade, também, de inserção de informações enciclopédicas no banco (figura 8). Embora a tela ofereça a possibilidade de inclusão de áudio e vídeo, essas facilidades ainda não foram implementadas na versão atual do programa. Estas informações foram retiradas da enciclopédia online Wikipédia (disponível em: http://pt.wikipedia.org/wiki/Página_principal), cujo conteúdo não sofre restrições de cópia:

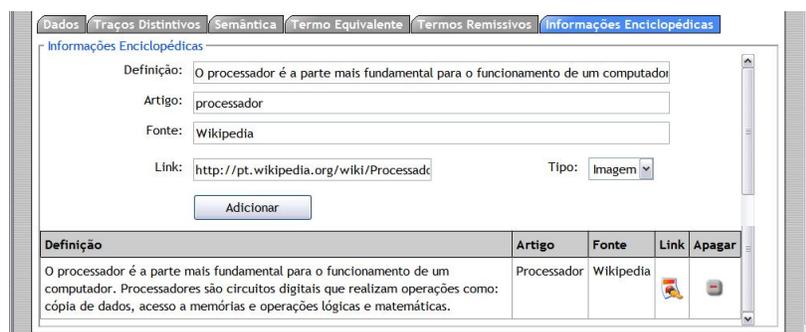


Fig. 8: Tela de inserção e visualização de informações enciclopédicas

9 O Enunciado Terminográfico resultante

Somando todas as possibilidades de paradigmas acima apresentadas, a construção do banco de dados oferece a seguinte fórmula para a construção do Enunciado Terminográfico de um termo:

**Termo = {+ entrada + enunciado terminográfico
(± PI + PD ± PS + PFE + PP ± Remissivas ± PE)}**

O sinal ± representa a opcionalidade (nem todos os campos podem ser preenchidos em virtude da carência de informações apresentadas pelos exemplos) e o sinal + representa a obrigatoriedade. A ordem dos componentes do enunciado terminográfico acima apresentado não representa, necessariamente, a ordem de visualização na página de consulta, já que a mesma apresenta diferentes possibilidades de construção da microestrutura. Pode-se tomar como exemplo o termo “processador”¹⁵:

processador. (CPU, UCP). *Hardware. s.m.s.* tipo de chip, parte principal do computador, também conhecido como CPU (Central Processing Unit); executa instruções (processos) e dados (através da UC - unidade de controle) e operações lógicas e aritméticas (através da UAL - unidade aritmética e lógica). *Ex.:* Processador (UCP). A UCP é a parte principal do computador responsável pelo processamento e execução de programas armazenados na memória principal. *Sinônimos:* CPU, UCP. *Hipônimo de:* computador. *Hiperônimo de:* UAL (unidade aritmética e lógica); UC (unidade de controle). *Co-hipônimos:* chip. *Veja Também:* chipset, CPU, UCP. **Córpus:** *Posição na Ordem de Frequência:* (67); *Nº de Ocorrências do termo:* (1212). **Informações Enciclopédicas:** O processador é a parte mais fundamental para o funcionamento de um computador. Processadores são circui-

¹⁵ Os campos que não puderam ser preenchidos não são disponibilizados na exibição da microestrutura.

tos digitais que realizam operações como: cópia de dados, acesso a memórias e operações lógicas e matemáticas. (*Processador – Wikipédia*)

processor. *Hardware. n.m/f.s.* flat square sliver of silicon device composed of circuits etched on its surface; it controls the computer and execute programs; also known as CPU (central processing unit). Ex.: The ‘brain’ of any computer is the central processing unit (CPU). This component does all the calculations and performs 90 percent of all the functions of a computer. There are many different types of processors for computers - so many, in fact, that you will learn about them later in this chapter in the section ‘Identifying Purposes and Characteristic of Processors’. *Synonyms:* CPU. *Hyponym of:* computer. *See Also:* CPU. **Corpus:** *Frequency order position:* (138); *Term number of occurrences:* (772). **Encyclopedic Information:** A central processing unit (CPU), or sometimes simply processor, is the component in a digital computer that interprets computer program instructions and processes data. CPUs provide the fundamental digital computer trait of programability, and are one of. (*CPU – Wikipedia*)

10 Considerações Finais

A confecção de trabalhos terminográficos ou lexicográficos em plataformas informatizadas proporciona maior padronização no resultado final, já que os pesquisadores devem seguir uma mesma linha de raciocínio para a construção da microestrutura. Embora a definição seja sempre um trabalho individual, no qual não pode haver um controle total do fazer, todos os outros paradigmas da microestrutura obedecerão à padronização pré-estabelecida.

O uso de corpora na construção de obras lexicográficas ou terminográficas já está se tornando uma realidade em várias partes do mundo. Muitos dicionários gerais de língua, especialmente os americanos e ingleses, usam corpora de grandes proporções, patrocinados pelas editoras ou governos, para elaborar suas obras.

TRADTERM, **15**, 2009, p. 133-154

Acredito que essa tendência seja irreversível, já que as estruturas (macro e micro) disponibilizadas nesses dicionários representam um estado atual da língua alvo, facilitando o entendimento por parte dos públicos leigo e especializado.

Com este artigo, espera-se um aprofundamento na análise das microestruturas em obras disponíveis, suas construções e possibilidades de padronização, tanto para tradutores quanto para outros públicos-alvo. Novos projetos podem ser adicionados ao banco de dados e à página já existentes; novos campos, além disso, podem vir a ser trabalhados no futuro, numa versão certamente aperfeiçoada desses componentes.

Referências bibliográficas

- ALMEIDA, G. M. B., PINO, D. H. P., SOUZA, D. S. L. (2007) A definição nos dicionários especializados: proposta metodológica. *RITerm – Debate Terminológico*, n° 3. Disponível em <http://www.riterm.net/revista/n_3/Art_Barcellos_Almeida.pdf>. Acesso em 15 de abril de 2007.
- ALVES, I. M. (1996) Definição terminológica: da teoria à prática. *TradTerm*. São Paulo: Humanitas, n° 3.
- ____ (2007) A renovação lexical nos domínios de especialidade. *Ciência e Cultura*, v. 58, n° 2. Disponível em <http://cienciaecultura.bvs.br/scielo.php?pid=S0009-67252006000200013&script=sci_arttext&tlng=pt>. Acesso em 23 de julho de 2007.
- AUBERT, F. H. (1996) *Introdução à metodologia da pesquisa terminológica bilíngüe*. São Paulo: Humanitas.
- BARBOSA, M. A. (1989) Da microestrutura dos vocabulários técnico-científicos. *Anais do IV Encontro Nacional da ANPOLL*. Recife: Anpoll.
- ____ (1994) Da microestrutura de vocabulários técnico-científicos bilíngües: para um microsistema terminológico de ecologia e meio ambiente. *IV Simpósio RITerm*. Disponível em <<http://www.riterm.net/actes/4simposio/barbosa.htm>>. Acesso em 11 de maio de 2007.
- ____ (1998) Terminologização, vocabularização, científicidade, banalização: relações. *Acta Semiotica et Linguística*, São Paulo: Plêiade, vol. 7.
- CABRÉ, M. T. (1993) *La terminología. Teoría, metodología, aplicaciones*. Barcelona: Editorial Antártida/Empúries, pp. 281-282.

- CAMPOS, M. G. (1994) *Sobre la elaboración de diccionarios monolingües de producción*. Granada: Universidad.
- FINATTO, M. J. B. (1994) Caracterização de paradigmas definicionais terminológicos. *IV Simpósio RITerm*. Disponível em <<http://www.riterm.net/actes/4simposio/bocorny.htm>>. Acesso em 11 de maio de 2007.
- ____ (1998) Microestrutura no dicionário terminológico: bases teóricas e orientações metodológicas em direção ao seu tratamento formal. *VI Simpósio RITerm*. Disponível em <<http://www.riterm.net/actes/6simposio/finato.htm>>. Acesso em 11 de maio de 2007.
- FRAÚSTO, I. P. (1990) Terminologia: uniformização vs. Criatividade. *II Simpósio RITerm*. Disponível em <<http://www.riterm.net/actes/2simposio/frausto.htm>>. Acesso em 11 de maio de 2007.
- FROMM, G. (2002) *Proposta para um modelo de glossário de informática para tradutores*. Dissertação de Mestrado em Linguística. São Paulo: Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo.
- ____ (2008) A construção e análise de *corpora* para alimentação de um banco de dados terminográfico: um exemplo. *Domínios de Linguagem*. Revista Eletrônica de Linguística (cf. www.dominiosdelinguagem.org.br), ano 2, nº 1.
- GONZÁLES-JOVER, A. G. & SIERRA, C. V. (2005) *Aspectos metodológicos para la elaboración de diccionarios especializados bilingües destinados al traductor*. Disponível em <http://www.ua.es/personal/chelo.vargas/Documentos/GomezYVargas_Toledo.pdf>. Acesso em 25 de abril de 2005.
- HAENSCH, G. (1982) Tipología de las obras lexicográficas e Aspectos prácticos de la elaboración de diccionarios. In: ETTINGER, S. et ali. *La lexicografía. De la lingüística teórica a la lexicografía práctica*. Madrid: Gredos.
- INTERNATIONAL STANDARD ORGANIZATION (1990). *ISO 1087*. Genebra.
- LARA, M. L. G. (2004) Diferenças conceituais sobre termos e definições e implicações na organização da linguagem documentária. *Ciência da Informação*. Brasília, vol. 33, nº 2, pp. 91-96. Disponível em <<http://www.scielo.br/pdf/ci/v33n2/a09v33n2.pdf>>. Acesso em 27 de julho de 2007.

- SAGER, J.C. (1990) *A practical course in terminological processing*.
Amsterdam: John Benjamin Publishing.
- TAGNIN, S. E. O. (2004) *Corpora: o que são e para quê servem*.
Minicurso. São Paulo.