

O processo de elaboração de um glossário bilíngue de colocações da área médica: contribuições para tradutores e professores de língua inglesa

Fernanda Goulart Ritti Dias
Universidade Federal de Alagoas
fergoulart@hotmail.com

Resumo

O objetivo deste estudo é apresentar o processo de elaboração de um glossário bilíngue (português-inglês) de colocações da área Médica, enfatizando a importância do conhecimento da terminologia adequada para o trabalho de tradução, revisão e ensino de língua inglesa. O estudo encontrou suporte teórico e metodológico na Linguística de *Corpus*, cujos principais pressupostos são a abordagem empírica na análise de dados e a visão probabilística da linguagem. Além disso, permite a investigação de grandes quantidades de dados por meio de ferramentas computacionais, por exemplo o *WordSmith Tools*. Para a execução do trabalho, foram compilados dois *corpora*, um em português e outro em inglês, com aproximadamente 85 mil palavras cada. Os resultados da pesquisa mostraram que o uso de *corpus* na tarefa de versão é muito relevante, principalmente quando se trata de colocações de uma área de especialidade. Concluímos que apesar de não ser uma tarefa simples e requerer muito cuidado e dedicação em todas as etapas, a elaboração de glossários é extremamente útil para aqueles que trabalham com a língua inglesa.

Palavras-chave: Linguística de *Corpus*; glossário bilíngue; colocações; língua inglesa.

Abstract

The objective of the present study is to show the process of organization of a bilingual collocations glossary (Portuguese-English) in the medical field, emphasizing the importance of knowing the correct terminology for translating, revising and teaching English. Theoretical and methodological support was provided by Corpus Linguistics, which emphasizes the importance of empirical data analyses and views language as a probabilistic system. Besides, it allows the investigation of large quantity of data through computational tools, e.g. *WordSmith Tools*. Two corpora, one in Portuguese and one in English, with approximately 85.000 words each were compiled.

The results indicated that the use of corpus in inverted translation tasks is very important, especially when it deals with collocations of a specific field. We concluded that although the organization of a glossary is not an easy task and requires great care and dedication in all its steps, it is very useful for those who work with the English language.

Keywords: Corpus Linguistics; bilingual glossary; collocations; English language.

1. Introdução

Grande parte das pesquisas científicas produzidas mundialmente é publicada em língua inglesa. Assim, cada vez mais estudantes e profissionais, além da necessidade de possuir habilidade de leitura nessa língua, precisam redigir seus trabalhos em inglês para conseguir publicação internacional e, dessa forma, obter prestígio acadêmico. Porém, o desconhecimento linguístico é uma grande barreira enfrentada pela maioria desses indivíduos, inclusive os tradutores que, muitas vezes, também não possuem familiaridade com a quantidade enorme de termos técnicos que cada área exige.

O desconhecimento de expressões apropriadas ao contexto por parte desses profissionais pode tornar o texto ininteligível, visto que muitas vezes acabam por fazer a tradução literal das palavras.

O fato é que cada área de especialidade utiliza sua própria linguagem, constituída por termos destinados a estabelecer maior precisão conceitual, nem sempre evidente para outras comunidades lingüísticas. (...) Ainda que os dicionários especializados tragam uma definição ou uma explicação, não parecem se preocupar com a dificuldade na escolha certa da combinação de palavras (CASTANHO, 2004, p. 3).

Os grupos profissionais são, assim, identificados por meio dos textos técnicos e científicos, expressos por uma linguagem que os marca, de um modo convencionalmente culturalmente estabelecido (KRIEGER e FINATTO, 2004). Dessa forma, com o intuito de atender à necessidade de estudantes, de professores e de tradutores de diversas áreas do conhecimento, algumas pesquisas sobre a elaboração de

glossários bilíngues com o apoio da Linguística de *Corpus* têm sido desenvolvidas no Brasil nos últimos anos.

Moraes (2005) elaborou uma proposta de glossário de colocações adverbiais (inglês-português) baseado em *corpus*. Carvalho (2007) apresentou uma metodologia para a elaboração de um glossário bilíngue com base em um *corpus* do domínio técnico-automobilístico. Costa Filho (2008) desenvolveu uma proposta de glossário bilíngue de termos utilizados na Teoria da Metáfora Conceitual. Navarro (2012) fez uma proposta de glossário bilíngue (inglês - português) de colocações da hotelaria, voltado às necessidades do tradutor. Seemann (2012) construiu um glossário bilíngue e bidirecional dos termos futebolísticos mais frequentes no par linguístico espanhol-português, usados rotineiramente na comunicação escrita.

Com o advento da Linguística de *Corpus* e o uso e exploração de *corpora* como metodologia de pesquisa, o levantamento de palavras e colocações se tornou maior e muito mais rápido. Os recursos eletrônicos disponíveis atualmente possibilitam a produção de glossários muito mais confiáveis em relação àqueles compilados antigamente, muitas vezes manualmente (ORENHA, 2004). De fato, examinar um *corpus* de um milhão de palavras pode parecer uma tarefa quase impossível para o ser humano, mas é algo que o computador faz em poucos segundos (BERBER SARDINHA, 1999).

Diante do exposto, o objetivo principal deste estudo é apresentar o processo de uma amostra de um glossário bilíngue (português-inglês) de colocações da área Médica, subárea Reumatologia, com enfoque em *fibromialgia*. Reumatologia é a área do conhecimento médico que se ocupa do estudo das doenças que acometem os tecidos conjuntivos. A fibromialgia é uma dessas doenças e refere-se a uma condição dolorosa generalizada e crônica.

Para a realização do nosso trabalho, dois *corpora*, um em português e um em inglês, cada um contendo aproximadamente 85 mil palavras, foram compilados em formato eletrônico e posteriormente analisados por meio do programa computacional denominado *WordSmith Tools*. Esse programa coloca à disposição do analista uma série de recursos, os quais, se bem usados, são extremamente úteis e poderosos na análise de vários aspectos da linguagem (BERBER SARDINHA, 1999).

Para maior confiabilidade e autenticidade dos dados, foram coletados apenas artigos científicos retirados de revistas especializadas e redigidos por profissionais da área nativos para ambas as línguas.

2. Pressupostos teóricos

A Linguística de *Corpus* é, segundo Teixeira (2004, p.34):

(...) uma metodologia de pesquisa específica em que a exploração da linguagem é feita por meio de evidências empíricas extraídas, por meio do uso de ferramentas computacionais, de um corpus de linguagem natural/autêntica, criteriosamente compilado e consultável eletronicamente.

Muitos pesquisadores dedicaram-se à descrição da linguagem por meio de *corpora* no início do século XX, entre eles educadores como Thorndike e linguistas, a exemplo de Boas e Fries. A grande diferença entre essa época e hoje é que os *corpora* não eram eletrônicos, ou seja, eram coletados, mantidos e analisados manualmente. Além disso, a ênfase era dada ao ensino de línguas, e não à descrição da linguagem, como atualmente, embora ainda haja o interesse no emprego de *corpora* na sala de aula e na investigação da linguagem de alunos de língua (BERBER SARDINHA, 2004).

A Linguística de *Corpus* requer o uso de um *corpus*. Segundo Bowker e Pearson (2002, p.9), “um *corpus* é simplesmente um corpo de texto”. No entanto, no contexto da Linguística de *Corpus*, “o *corpus* pode ser descrito como uma ampla coletânea de textos autênticos coletados em formato eletrônico de acordo com um critério específico”. Os textos devem ser selecionados de acordo com um critério explícito para que possam ser utilizados como uma amostra representativa de uma língua em particular. O critério utilizado para elaborar um *corpus* dependerá do propósito do estudo” (id. *ibid.*, p. 10).

Outra importante área do conhecimento utilizada para a elaboração de glossários é a Terminologia, que “diz respeito à investigação das propriedades linguísticas, conceituais e pragmáticas

das unidades terminológicas com vistas à produção de dicionários técnicos e científicos” (BESSÉ ET ALII, apud Krieger e Finatto, 2004, p.130).

Ao apoiar-se cada vez mais na Linguística, a Terminologia tem oferecido fundamentos teóricos que auxiliam no reconhecimento dos termos e na análise de seu comportamento nas diferentes línguas. Com isso, oferece aos tradutores uma compreensão mais clara sobre a natureza, o estatuto, a constituição e o funcionamento dos termos técnico-científicos (KRIEGER e FINATTO, 2004).

A Terminologia tem como campo de atuação as línguas de especialidade e a sua unidade padrão é o termo (BARROS, 2004). A identificação de um termo costuma ser problemática para os tradutores, uma vez que sua estruturação apresenta unidades lexicais complexas. Além disso, não é fácil demarcar as fronteiras rígidas entre o léxico especializado e o geral (KRIEGER e FINATTO, 2004).

Em um dicionário de língua geral, encontramos somente o vocabulário “técnico” de base de grandes áreas especializadas, ao passo que em um dicionário terminológico são empregados os termos mais específicos da área.

2.1. As colocações

O termo colocação foi originalmente introduzido pelo linguista J. R. Firth para descrever o fato de algumas palavras “andarem juntas”, produzindo um discurso natural, tanto falado quanto escrito (MORAES, 2005).

Segundo Tagnin (2005), não há, via de regra, explicação para o fato de certas palavras combinarem-se com tal naturalidade. É o uso que consagra a maneira a serem utilizadas na língua. Um exemplo básico e simples citado pela autora é *cão e gato*. Por que não ocorre *cachorro e gato*? Simplesmente porque o uso consagrou *cão e gato* e não *cachorro e gato* nessa colocação.

Às vezes, só saberemos que uma combinação é uma colocação quando, ao procurarmos certa palavra em diversos dicionários, descobriremos que todos eles, ou pelo menos a grande maioria, apresenta a mesma combinação. Em alguns casos pode ocorrer mais de uma opção, mas sempre dentro do mesmo campo semântico. Em português temos o caso de ‘coroca’, que, preferencialmente ocorre com

O processo de elaboração de um glossário bilíngue

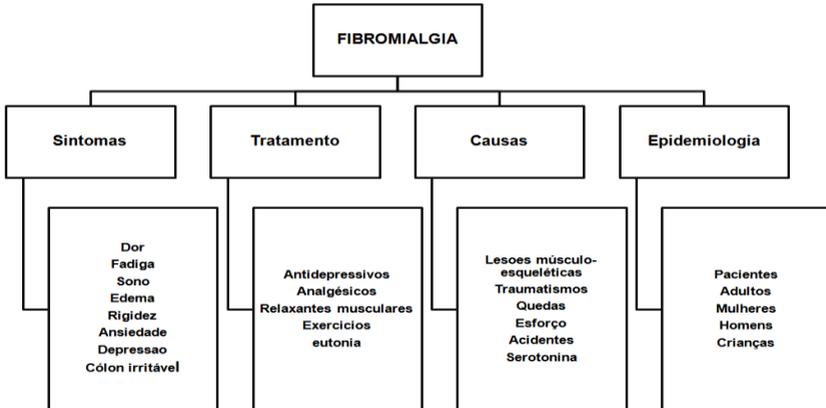
‘velha’ (velha coroca); e de ‘varrido’, que ocorre com ‘louco’ ou ‘doido’ (louco/doido varrido) (id. *ibid.*, p. 37-38).

As categorias de colocações expostas pela autora são: *colocações adjetivas* (amigo íntimo), *colocações nominais* (cartão de crédito), *colocações verbais* (marcar um encontro) e *colocações adverbiais* (amar cegamente).

3. Metodologia da pesquisa

Primeiramente, construímos uma árvore de domínio, ou seja, uma organização semântica da área-objeto a fim de delimitar o conjunto terminológico e determinar a pertinência dos termos. Na Figura 1, apresentamos uma amostra da árvore de domínio realizada na nossa pesquisa.

Figura 1. Amostra da árvore de domínio



Dividimos nosso trabalho em seis etapas:

1. Busca dos *corpora* na *web*.
2. Colocação dos cabeçalhos e etiquetas necessárias.
3. Produção da lista de palavras-chave de ambos os *corpora*, (com a utilização do programa *WordSmith*)

Tools) obtida por meio da comparação da frequência da lista de palavras do *corpus* de estudo com a frequência da lista de palavras dos *corpus* de referência.

4. Levantamento das palavras-chave de conteúdo de cada lista.
5. Extração das colocações por meio da ferramenta *Concord* do *WordSmith Tools* e pela lista de *Clusters*.
6. Compilação do glossário.

3.1. Coleta e armazenamento dos *corpora*

A metodologia empregada na pesquisa constituiu-se da coleta de dois *corpora*: um em português e outro em inglês, cada um contendo, aproximadamente, 85 mil palavras. Em vista dos objetivos deste estudo, selecionamos textos autênticos, redigidos por especialistas da área Médica, retirados de *sites* confiáveis e consagrados da área.

A seleção dos textos seguiu os seguintes critérios de coleta:

1. Somente artigos científicos originais completos que se tratavam do tema fibromialgia, produzidos por falantes nativos de cada uma das línguas escolhidas.
2. Textos disponíveis eletronicamente em *sites* confiáveis. A busca dos textos na língua inglesa foi feita somente no *site* www.pubmed.com, um *site* em inglês que ajuda a realizar pesquisas bibliográficas por meio do acesso à base bibliográfica *MedLine*, desenvolvida pela NLM (National Library of Medicine) e que conta com mais de 12 milhões de citações. Já a busca dos textos em português foi feita no *site* www.scielo.br, uma biblioteca eletrônica que abrange uma coleção selecionada de periódicos científicos brasileiros.
3. Para uma boa análise, optamos por conseguir um número aproximado de palavras em cada *corpora*, o que, conforme citado anteriormente, totaliza 85 mil.
4. O ano de publicação dos textos também recebeu atenção durante a coleta. Escolhemos textos

publicados somente entre 2000 e 2006. A comparabilidade dos corpora é apresentada na tabela 1.

Tabela 1. Características gerais dos corpora comparáveis

	Língua Portuguesa	Língua Inglesa
Tamanho	84.708	85.199 palavras
Número de textos	21	20
Meio	Escritos	Escritos
Fonte	www.scielo.br	www.pubmed.com
Assunto	Fibromialgia	Fibromialgia
Gênero textual	Artigos científicos	Artigos científicos completos
Autoria	Autores nativos	Autores nativos
Língua	Português	Inglês
Data de publicação	De 2000 a 2006	De 2000 a 2006

Os textos encontrados foram separados por língua, colocados em arquivos no computador e nomeados com um código, (por exemplo, FIB1P, quando em português; e FIB1E, quando em inglês.). Para armazená-los, utilizamos o Bloco de Notas do *Windows*, que grava os textos automaticamente em *.txt*. Esse tipo de arquivo é requerido pelo *Wordsmith Tools*, o qual, por sua vez, foi nossa ferramenta de análise do *corpus*.

Em cada texto foi inserido um cabeçalho, que é parte do arquivo de cada texto do *corpus* que contém informações sobre o texto, tais como a origem, a data de coleta, a autoria, entre outras.

3.2. Análise dos corpora

Para a construção do nosso trabalho, utilizamos a “versão 3” do programa *WordSmith Tools*, disponível no site <http://www.lexically.net/wordsmith/>.

As ferramentas do programa utilizadas foram as seguintes: 1- *Wordlist* (lista de palavras individuais); 2- *KeyWords* (lista de palavras-chave) e 3- *Concord* (concordância; lista de agrupamentos lexicais-clusters. Primeiramente, fizemos uma lista de palavras, e, em seguida, a lista de palavras-chave de cada *corpora* de estudo.

Para a pesquisa e montagem da lista de palavras-chave do nosso *corpus*, utilizamos dois *corpora* de língua geral como referência: em inglês, o *Brown Corpus*¹, que possui textos escritos em inglês americano e possui um milhão de palavras. Para os textos em português, utilizamos o *Corpus Lacio-Ref*, que disponibiliza um conjunto de textos escritos autênticos do português brasileiro e possui 4.156.816 palavras.

Com base na lista de palavras-chave do *corpus* de línguas portuguesa, selecionamos as palavras que ocorreram com maior frequência e iniciamos a busca de colocações. Com a ferramenta *Concord* do *WordSmith Tools*, obtivemos as listas de concordâncias (Figura 2). Encontradas as concordâncias na língua portuguesa, passamos a buscar as concordâncias equivalentes na língua inglesa.

Encontradas as concordâncias, iniciamos a organização do glossário, que será apresentada na próxima seção.

4. Resultados e discussão

Com base na lista de palavras-chave do *corpus* em língua portuguesa, foram gerados 21 termos-chave: *ansiedade, depressão, diagnóstico, distúrbio, dor, eutonia, exercício, fadiga, fibromialgia, hepatite, melhora, mulher, paciente, percepção, prevalência, qualidade, síndrome, sintoma, sintomatologia, sono e tratamento*.

A partir desses 21 substantivos, foram geradas 53 colocações. Ressaltamos que foram selecionadas as colocações mais frequentes do *corpus* de estudo, e não um tipo específico de colocação, ou seja, verbal, nominal, etc.

O objetivo era encontrar todos os termos dentro do *corpus* de estudo, porém, isso nem sempre foi possível. Durante a busca das colocações, não encontramos três equivalentes na língua inglesa (*Efeito da eutonia, Prevalência de dor crônica difusa e Síndrome do cólon irritável*). Diante de tal limitação, optamos por buscá-las no site www.google.com.br, de modo a verificar a origem e a autenticidade dos textos sempre, além de garantirmos se o termo pertencia a textos

¹ Outro *corpora* de referência em língua inglesa que tem sido bastante utilizado por pesquisadores da área é o *British National Corpus* (BNC).

O processo de elaboração de um glossário bilíngue

sobre fibromialgia. As três colocações buscadas na *internet* foram encontradas no *site PUBMED*, utilizado no início da pesquisa, o que garantiu mais segurança quanto às colocações encontradas.

Figura 2. Exemplo de lista de concordância com termo de busca “pacientes”



N	Concordance
1	quando $p < 0,05$. RESULTADOS Entre os 166 pacientes portadores de FM, 162 (97,6%) era
2	com o valor de $p < 0,05$. Foram avaliados 84 pacientes , 50 homens (59,52%) e 34 mulheres
3	no último ano foi 1 ($\pm 0,5$) e somente 4 (3,3%) pacientes apresentaram mais de 2 parceiros s
4	2,0 0,99 por recomendação de outros pacientes . O atendimento foi custeado pelo pr
5) 0,5 tabela 1 dados demoGRáficos dos 222 pacientes paRticipantes da pesquisa sintomas
6	quadRo 1 oRGanoGRAMa dos Resultados de pacientes com HipotiReoidismo Paciente com
7	s anticorpos anti-Coxsackie tipos 1 a 6 em 548 pacientes com fadiga crônica, divididos em su
8	a média de 1,4cm, considerando- se as quatro pacientes , resultando em uma redução de 2,4
9	rograma de reabilitação 10. Cerca de 75% dos pacientes com fibromialgia queixam-se de um
10	po pelo menos 11 tender points. Foram excluídas pacientes portadoras de patologias sistêmicas
11	l, v. 46, n. 1, p. 11-15, jan/fev, 2006 Pacientes Pacientes Não houve diferença estatística qua
12	ntrada em 113 (94,2%) pacientes ; 105 (87,5%) pacientes se queixaram de ansiedade e 47 (39,
13	onto sensível. Após a 11a e última sessão, as pacientes foram novamente reavaliadas individu
14	ro de tender-points menor que 14; ** Grupo 2 = pacientes com número de tender-points igual o
15	bromialgia, com idades entre 18 e 45 anos. As pacientes e os controles foram questionados q
16	deviam ter mais do que 18 anos de idade. Os pacientes não poderiam estar em uso de opiói
17	drome, em agudo (19 pacientes) e crônico (20 pacientes). Anticorpos IgM contra enterovirose
18	ção pelo parvovírus B 19 nesse grupo seletor de pacientes (28). VÍRUS COXSACKIE B O gêner
19	l, 1999; WINFIELD, 1999). Cerca de 20% dos pacientes com fibromialgia apresentam depres
20	o de FM do ACR. (2) critérios de exclusão As pacientes que não completaram, no mínimo, o
21	resentação(1, 2, 3). O acompanhamento dos pacientes baseia-se na evolução da sua sinto
22	1+ Outros 5 (25%) 5 (25%) A idade média das pacientes do grupo fibromialgia foi 34,9 anos e
23	e, anos (entre 20 e anos) sCI 3 (%) tabela 3 pacientes com fibRomialGia (n=200) PadrÃO n
24	abelas 2 e 3)Avaliação da satisfação sexual em pacientes portadores de fibromialgia Tabela 2.
25	balho em casa" na questão 3, uma vez que há pacientes que somente desenvolvem atividades

No quadro 1, apresentamos as colocações que apresentam preposições, o que parece ser uma das maiores dificuldades para quem trabalha com produção textual em língua estrangeira. As preposições *da*, *do* e *de*, em português, que têm a tendência de serem traduzidas como *of* em inglês, foram aqui apresentadas como *of*, *in*, ou até sem a preposição.

Quadro 1. Colocações com a preposição “de”

Colocações com a preposição of
Diagnóstico de fibromialgia – <i>Diagnosis of fibromyalgia</i>
Distúrbio de dor – <i>Disturbance of pain</i>
Presença de dor – <i>Presence of pain</i>
Prevalência de depressão – <i>Prevalence of depression</i>
Prevalência de dor crônica difusa – <i>Prevalence of chronic widespread pain</i>
Qualidade de vida – <i>Quality of life</i>
Qualidade do sono – <i>Quality of sleep</i>
Tratamento da fibromialgia – <i>Treatment of fibromyalgia</i>
Colocações com a preposição in
Alterações do sono – <i>Alterations in sleep</i>
Melhora da dor – <i>Improvement in pain</i>
Melhora da qualidade de vida – <i>Improvement in quality of life</i>
Redução da sintomatologia – <i>Reduction in symptoms</i>
Colocações sem preposição
Distúrbio do sono – <i>Sleep disturbance</i>
Efeito da eutonia – <i>Eutonia effect</i>
Escores de depressão – <i>Depression scores</i>
Intensidade da dor – <i>Pain intensity</i>
Melhora da sintomatologia – <i>Symptom improvement</i>
Pacientes do sexo feminino – <i>Female patient</i>
Percepção da dor – <i>Pain perception</i>
Programas de exercícios – <i>Exercise program</i>
Síndrome da fadiga crônica – <i>Chronic fatigue syndrome</i>
Síndrome da fibromialgia – <i>Fibromyalgia syndrome</i>
Síndrome do cólon irritável – <i>Irritable bowel syndrome</i>
Sintomas da dor – <i>Pain symptoms</i>
Sintomas da fibromialgia – <i>fibromyalgia symptoms</i>
Vírus da hepatite C – <i>Hepatitis C virus</i>

Também encontramos a preposição *com*, com equivalentes a exemplo de *with* ou simplesmente sem a preposição (quadro 2).

Diante disso, não restam dúvidas quanto à importância do levantamento das colocações em outra língua por meio de *corpora* eletrônico. Ficou claro que, quando se trata de colocações, deve-se ter cuidado dobrado, pois nem tudo o que parece é. Assim, usuários de

O processo de elaboração de um glossário bilíngue

uma determinada língua estrangeira, neste caso a língua inglesa, devem buscar obras que atendam à busca por soluções tradutórias para as colocações, haja vista que a maioria dos dicionários técnicos e especializados existentes não apresentam a forma consagrada que essas colocações se apresentam. De fato, segundo Tagnin (2005), às vezes só percebemos que uma combinação é uma colocação quando vários dicionários apresentam determinadas palavras da mesma forma.

Quadro 2. Colocações com a preposição “com”

Colocações com a preposição <i>with</i>
Mulheres com fibromialgia – <i>Women with fibromyalgia</i>
Pacientes com fibromialgia – <i>Patients with fibromyalgia</i>
Pacientes com síndrome da fibromialgia – <i>Patients with fibromyalgia syndrome</i>
Colocação sem preposição
Pacientes com dor crônica – <i>Chronic pain patients</i>

A preposição *na* também merece destaque, já que teve por equivalentes *on* e *in* (quadro 3).

Quadro 3. Colocações com a preposição “na”

Colocações com as preposições <i>on</i> e <i>in</i>
Impacto na qualidade de vida – <i>Impact on quality of life</i>
Melhora na qualidade de vida – <i>Improvement in quality of life</i>

É possível, pois, afirmar que a consciência da convencionalidade de uma língua é fator crucial para que uma comunicação ou tradução seja fluente e efetiva. A falta dessa consciência pode, sem dúvida, “comprometer transações de negócios, pelo simples motivo de não seguir regras composicionais e convencionais ditadas pela língua” (ORENHA, 2004, p. 2).

4.1. A organização do glossário

Ressaltamos novamente que nosso propósito foi mostrar como se dá o processo de elaboração de uma amostra glossário técnico de

colocações. Dessa forma, apenas uma pequena amostra do glossário será apresentada.

Os termos retirados da lista de palavras-chave, ou seja, os substantivos formaram as bases das colocações, e, portanto, foram dispostos em ordem alfabética e no singular. Além disso, foram destacados e sublinhados para facilitar a busca do consulente. Dentro de cada base, foram apresentadas as entradas, ou seja, as colocações que fazem referência ao respectivo substantivo de base, as quais também foram organizadas alfabeticamente.

Embora não tenhamos apresentado, em nossa pequena amostra, informações etimológicas, ortográficas, fonéticas ou gramaticais sobre as entradas, sabemos de sua importância para aqueles que trabalham com a produção textual. Definições ou explicações sobre as entradas também não foram fornecidas em nossa amostra. De fato, segundo Barros (2004, p.145), a principal característica de um glossário “é não apresentar definições, mas tão-somente uma lista de unidades lexicais ou terminológicas acompanhadas de seus equivalentes em outras línguas”.

Os equivalentes na língua de chegada seguem as colocações referentes ao substantivo base e podem ser lidos horizontalmente. Abaixo das colocações foram fornecidos exemplos em ambas línguas, e não a tradução dos mesmos, retirados do nosso *corpus* de estudo com a ajuda da ferramenta *concord*. Em cada exemplo, as colocações foram destacadas para melhor visualização.

A tabela 2 apresenta uma pequena amostra da organização do glossário. Selecionamos alguns exemplos da letra “D”.

5. Considerações finais

O computador e a linguística, aliados, têm trazido inúmeros benefícios para estudiosos da área. Inúmeras tarefas, de diversas áreas do saber, antes impossíveis ou difíceis de serem realizadas, tornaram-se possíveis e mais rápidas. A ferramenta computacional *WordSmith Tools*, dotada de inúmero recursos, realiza etapas essenciais para a compilação de glossários e dicionários, que jamais poderiam ser feitas manualmente com tanta eficácia, precisão e em tão pouco tempo.

Tabela 2. Amostra de organização do glossário

D

DEPRESSÃO

ANSIEDADE E DEPRESSÃO

A influência do estilo de vida no desenvolvimento de **ansiedade e depressão** é comprovada.

ESCORES DE DEPRESSÃO

Relação entre a percepção da saúde versus os **escores de depressão**.

PREVALÊNCIA DE DEPRESSÃO

Neste estudo, buscou-se descrever a relação entre a **prevalência de depressão** e sua relação com a qualidade de vida nos pacientes com fibromialgia.

DEPRESSION

ANXIETY AND DEPRESSION

Anxiety and depression are associated with the reporting of greater subjective pain.

DEPRESSION SCORES

Unexpectedly, serum serotonin levels rose correspondingly with depression scores, contrary to what has been reported in clinic patients.

PREVALENCE OF DEPRESSION

Finally, an increased prevalence of depression has been found in patients with fibromyalgia.

É claro que a contribuição do linguista é de extrema importância no processo, pois é ele quem seleciona os textos a serem inseridos no *corpus* e analisa os termos que serão inclusos na obra de referência. Ressaltamos que se essas etapas não forem executadas com muito critério, podem, sem dúvida, afetar o produto final.

A experiência obtida durante a compilação de uma amostra de glossário bilíngue ratifica essas afirmações. A coleta, o armazenamento dos textos e a seleção das colocações que compuseram nosso glossário foram as etapas mais trabalhosas e que mereceram mais atenção. No entanto, sem a ajuda do computador, em especial da *internet*, que

disponibiliza textos autênticos e de diversas áreas do saber, o trabalho não poderia ter sido iniciado.

Em geral, o principal objetivo do trabalho, descrever o processo que compõe a elaboração de um glossário bilíngue com base na Linguística de *Corpus*, foi realizado. Também foi possível mostrar as peculiaridades das colocações e, conseqüentemente, sua importância para os que trabalham com produção de texto em língua inglesa. Ficou claro que o uso de *corpus* na tarefa de versão é muito relevante, principalmente quando se trata de colocações de uma área de especialidade. Além disso, quanto maior o *corpus* de estudo, mais eficaz se torna a busca das colocações. Por isso, a tarefa de elaboração de um glossário não é tarefa fácil e requer muito cuidado e dedicação em todas as etapas.

Referências

BARROS, Lidia Almeida. *Curso básico de terminologia*. Edusp, 2004.

BERBER SARDINHA, Tony. Usando WordSmith Tools na investigação da linguagem. *DIRECT papers* 40, 1999. Disponível em: http://www2.lael.pucsp.br/direct/direct_papers.htm. Acesso em 10 dez. 2005.

_____. A influência do tamanho do corpus de referência na obtenção de palavras-chave. *DIRECT papers*, 38, 1999. Disponível em: http://www2.lael.pucsp.br/direct/direct_papers.htm. Acesso em 8 dez. 2005.

_____. O banco de palavras chave. *LAEL, PUCSP. DIRECT Papers*, 39, 1999. Disponível em: http://www2.lael.pucsp.br/direct/direct_papers.htm. Acesso em 8 dez. 2005.

_____. *Linguística de Corpus*. Barueri, SP: Manole, 2004.

BOWKER, Lynne; PEARSON, Jennifer. *Working with specialized language*. London/New York: Routledge, 2002.

CARVALHO, Estela Maria Faustino. *Metodologia de construção de um Glossário bilíngüe com base em um corpus de Domínio técnico*. 2007. 80f. Dissertação (Mestrado em Estudos da Tradução). Centro de Comunicação e Expressão. Universidade Federal de Santa Catarina, Florianópolis, 2007.

CASTANHO, Rosa Maria. *Proposta para a Elaboração de um Glossário de Colocações na Área Médica. Subárea Hipertensão Arterial*. 2004. 92p. – *Dissertação*. (Mestrado em Estudos Linguísticos e Literários em Inglês) – Departamento de Letras Modernas da Faculdade de Filosofia, Letras e Ciências Humanas da Universidade de São Paulo, São Paulo, 2004.

COSTA FILHO, José Edelberto. *Elementos para um glossário bilíngüe (português e inglês) de termos-chave da Teoria da Metáfora Conceitual*. 2008. 148f. Dissertação (Mestrado em Linguística Aplicada). Departamento de Letras. Universidade Estadual do Ceará, Fortaleza, 2008.

KRIEGER, Maria da Graça; FINATTO, Maria José Bocorny. *Introdução à terminologia*. Ed. Contexto, 2004.

MORAES, Helmara Febeliana Real de. *O tradutor pode estar redondamente enganado: um estudo contrastivo de colocações adverbiais (inglês-português) sob o enfoque da linguística de corpus*. 2005. Dissertação. 207p. (Mestrado em Letras). Faculdade de Filosofia, Letras e Ciências Humanas da Universidade de São Paulo, São Paulo, 2005.

NAVARRO, Sandra Lago Martinez. *Glossário bilíngüe de colocações de hotelaria: um modelo à luz da Linguística de Corpus*. 2011. Dissertação (Mestrado em Estudos Linguísticos e Literários em Inglês) - Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo, 2012.

ORENHA, Adriane. *A compilação de um glossário bilíngüe de colocações, na área de Jornalismo de Negócios, baseado em corpus comparável*. 2004. 246 f. Dissertação (Mestrado em Estudos

Linguísticos e Literários), Faculdade de Letras e Ciências Humanas, Universidade de São Paulo, São Paulo, 2004.

_____. Aplicações léxico-terminográficas da linguística de corpus: relato da elaboração de um glossário bilíngue de colocações na área de negócios. *Intercâmbio*, v. 13, n. 1, 2004. Disponível em: <http://revistas.pucsp.br/index.php/intercambio/article/view/3992>. Acesso em 20 jan. 2007.

SEEMANN, Paulo Augusto Almeida. *A construção de um glossário bilíngue de futebol com o apoio da Linguística de Corpus*. 2012. 228 f. Dissertação (Mestrado em Língua Espanhola e Literaturas Espanhola e Hispano-Americana) - Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo, 2012.

TAGNIN, Stella Esther Ortweiler. *O jeito que a gente diz: expressões convencionais e idiomáticas*. São Paulo: Disal, 2005.

TEIXEIRA, Elisa Duarte. *Receita qualquer um traduz. Será? A Culinária como área técnica de tradução*. 2004. Dissertação (Mestrado em Estudos Linguísticos e Literários em Inglês) –Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo, 2v, 2004.

Recebido em: 16/09/2011

Aceito em: 01/10/2012

Title: The process of organization of a bilingual collocations glossary in the medical field: contributions for translators and English teachers

