

**UMA REFLEXÃO SOBRE A ELABORAÇÃO DE UM LÉXICO  
BILÍNGÜE PRELIMINAR NA SUBÁREA DE CARDIOLOGIA A  
PARTIR DO USO DE TERMOS ENCONTRADOS EM UM CORPUS  
PARALELO E EM DOIS CORPORA COMPARÁVEIS**

(A Reflection on the Elaboration of a Preliminary Bilingual Lexicon of  
Cardiology Based on the Terms Used in a Parallel and in Two  
Comparable Corpora)

Paula Tavares Pinto PAIVA  
Profa. Dra. Diva Cardoso de CAMARGO  
Profa. Dra. Cláudia Maria XATARA  
(Unesp – Universidade Estadual Paulista – S.J.R.P.)

*RESUMO: O objetivo principal deste trabalho foi propor uma reflexão sobre o processo a ser utilizado para a elaboração de um léxico bilíngüe na subárea de cardiologia. Para tanto, tomamos como base os conceitos dos estudos da tradução baseados em corpus, da lingüística de corpus e da terminologia. Como material para compor os corpora utilizamos artigos de cardiologia escritos em português e traduzidos para o inglês, assim como artigos originalmente escritos em português e em inglês. Com base no léxico proposto, pudemos notar algumas diferenças e algumas correspondências de uso entre os termos que aparecem no subcorpus de estudo de textos originais e traduzidos e nos corpora comparáveis em português e em inglês. Essa diferença apontaria que os termos não seriam unívocos dentro dessa linguagem de especialidade devido às diferenças de uso pelos especialistas de cardiologia para designar um mesmo referente.*

*PALAVRAS-CHAVE: termos médicos em cardiologia; estudos da tradução baseados em corpus; lingüística de corpus; terminologia.*

*ABSTRACT: This paper aims at proposing a reflection on the process for elaborating a bilingual lexicon of cardiology. The methodology used is that of corpus-based translation studies; that of corpus linguistics; as well as that of terminology. The material used in this study was composed of papers of cardiology written in Portuguese and translated into English and articles originally written in Portuguese and in English. Based on the lexicon proposed we have seen that there are some differences and correspondences between*

*the use of terms in the articles from the subcorpus of study composed of original and translated texts and the use of terms in the comparable corpora in Portuguese and in English. This difference indicates that the terms are not univocal in the medical language due to the differences in use by cardiologists to name the same referents.*

**KEY-WORDS:** *medical terms in cardiology; corpus-based translation studies; corpus linguistics; terminology.*

## 0. Introdução

A medicina é um campo que sempre recebeu atenção especial de toda a sociedade por se tratar de um assunto de primordial importância na vida do homem – a saúde humana. Nos estudos lingüísticos, em particular nas pesquisas em tradução, lexicologia e terminologia, têm-se observado trabalhos que datam de épocas as quais levantamentos de termos estão ligados a estudos precursores no campo da biologia.

Segundo Barros (2004:29), “já no primeiro século da era cristã o gramático Herodianus e o médico Heródoto elaboraram glossários que explicavam os termos médicos utilizados pelo grego Hipócrates (c. 460-377 a.C.), o pioneiro na descrição sistemática do corpo humano”. A autora também destaca trabalhos mais recentes realizados dentro de línguas de especialidade, como a proposta de termos novos ou neônimos, para designar conceitos de uma obra estrangeira especializada na língua de chegada.

Sabemos que, nos dias de hoje, a medicina continua sendo uma área fértil na literatura mundial e que merece ser tratada cuidadosamente por todos os profissionais que estão, de alguma maneira, ligados a ela, como os tradutores, terminólogos, lexicógrafos e demais pesquisadores.

Nesse artigo, apresentaremos uma reflexão sobre a elaboração de um léxico bilíngüe composto por termos médicos em português e em inglês, realizado a partir de revistas bilíngües e revistas escritas originalmente em inglês e em português. Esse léxico é importante por observar o conjunto léxico utilizado por autores e tradutores ao veicular as pesquisas brasileiras divulgadas no exterior. Para tanto, propomos a compilação de um léxico bilíngüe de cardiologia voltado para o tradutor, tomando como base conceitos da tradução, da lingüística de corpus da terminologia e da lexicologia. Tal compilação, feita do ponto de vista do tradutor, poderá ser

continuada e aprofundada por estudiosos da lexicologia/lexicografia e da terminologia/terminografia.

Para uma definição de léxico, nesse trabalho, fundamentamo-nos em Boutin-Quesnel *et al.* (1985:30), para quem o léxico é um “repertório que registra termos acompanhados de seus equivalentes em uma ou mais línguas, e que não apresenta definições. Nota: Os léxicos, em geral, abrangem um só domínio”<sup>1</sup>.

O léxico comentado no presente trabalho será compilado com base em um corpus paralelo constituído de quinze artigos escritos originalmente em português e as respectivas traduções para o inglês. Por corpora paralelos, entendemos “corpora de textos fonte e suas respectivas traduções” (Baker, 1993:248). Além do corpus principal (paralelo), utilizaremos dois corpora de controle (comparáveis), um deles composto por artigos escritos originalmente em português e o outro por artigos originalmente escritos em inglês.

## 1. Fundamentação Teórica

No caso da pesquisa efetuada para a elaboração do léxico, é o discurso terminológico que estará em foco, dada a especificidade dos termos de cardiologia. Acreditamos, contudo, que seja importante esclarecer aqui o campo de atuação da terminologia. Andrade, ao apresentar suas particularidades, mostra-nos uma definição bastante apropriada:

A terminologia pode ser encarada como uma “especificidade” da lexicologia, uma vez que trata, não de todas as palavras da língua, mas daquelas que constituem as linguagens especializadas [...] entre suas diversas tarefas cabe-lhe o estudo das relações de significação (expressão e conteúdo) do signo terminológico, o que inclui a complexa dinâmica da criação desse signo (neonímia), e da renovação e ampliação dos universos de discursos terminológicos. (Andrade, 2001:192).

A autora também destaca que “quanto ao objeto, [...] a terminologia se ocupa do termo, ou seja, da palavra especializada, dos conceitos inerentes às diversas matérias especializadas” (Andrade, 2001:192).

---

1. *Répertoire, qui inventorie des termes accompagnés de leurs équivalents dans une ou plusieurs autres langues, et que ne comporte pas de définitions. Note.- Les lexiques portent généralement sur un seul domaine.* [A tradução das citações é de nossa responsabilidade].

2. *Parallel corpora, that is corpora of source texts and their translations.*

Também a lingüística de corpus tem função primordial para a tradução. Em virtude de trabalhar com dados reais utilizando corpora eletrônicos, a lingüística de corpus tem fornecido uma abordagem que possibilita a observação de redes semânticas e campos lexicais, o que facilita ao analista trabalhar com grandes quantidades de dados.

Segundo Beber Sardinha, a definição que engloba as características principais de um corpus computadorizado é a de Sanchez:

Um conjunto de dados lingüísticos (pertencentes ao uso oral ou escrito da língua, ou a ambos), sistematizados segundo determinados critérios, suficientemente extensos em amplitude e profundidade, de maneira que sejam representativos da totalidade do uso lingüístico ou de algum de seus âmbitos, dispostos de tal modo que possam ser processados por computador, com a finalidade de propiciar resultados vários e úteis para a descrição e análise. (Sanchez, 1995:8-9, *apud* Berber Sardinha, 2000:338).

Essa definição, para o autor, reúne os mais importantes pontos a serem seguidos na construção de um corpus, a saber:

- (a) A origem: os dados devem ser autênticos.
- (b) O propósito: o corpus deve ter a finalidade de ser um objeto de estudo lingüístico.
- (c) A composição: o conteúdo do corpus deve ser criteriosamente escolhido.
- (d) A formação: os dados do corpus devem ser legíveis por computador.
- (e) A representatividade: o corpus deve ser representativo de uma língua ou variedade.
- (f) A extensão: o corpus deve ser vasto para ser representativo. (Berber Sardinha, 2004:18-19).

A representatividade e a extensão do corpus estão diretamente ligadas às necessidades do pesquisador. A extensão do corpus comporta três dimensões: o número de palavras (quanto maior o número de palavras maior será a possibilidade de o corpus conter palavras de baixa frequência); o número de textos do corpus específico (o número maior de textos garante que o tipo textual, gênero ou registro, esteja mais adequadamente representado); número de gêneros, registros ou tipos textuais (no caso de corpora variados, criados para representar uma língua como um todo). No léxico proposto, embora os subcorpora compilados se enquadrem na especificação de tamanho pequeno-médio, como veremos abaixo, consideramos os corpora a serem utilizados como representativos do gênero analisado, ou

seja, artigos publicados em revistas bilíngües de uma subárea médica, a de cardiologia. São representativos também por apresentar textos que seguem o mesmo padrão lingüístico exigido por revistas conceituadas na área médica.

Já a questão do tamanho do corpus é de natureza subjetiva, ou seja, “os usuários de um corpus atribuem a ele a função de serem representativos de uma certa variedade”; também cabe a esses usuários “serem cuidadosos em relação à generalização dos seus achados para uma população inteira (um gênero ou a língua inteira, por exemplo)” (Beber Sardinha, 2000:345).

A seguinte classificação em relação ao tamanho de um corpus é feita pelo autor:

Tabela 1: Classificação do tamanho do corpus

<i>Tamanho em palavras</i>	<i>Classificação</i>
Menos de 80 mil	Pequeno
80 a 250 mil	Pequeno-médio
250 mil a 1 milhão	Médio
1 milhão a 10 milhões	Médio-grande
10 milhões ou mais	Grande

Desse modo, o corpus paralelo de estudo, composto por 15 artigos médicos e suas respectivas traduções para o inglês, pode ser classificado como de porte pequeno, visto que é composto por 54.613 palavras nos TTs e nos TOs por 53.220. Já os corpora comparáveis, constituídos por textos originalmente escritos nas línguas de partida e de chegada, se enquadrariam nos corpora de classificação pequeno-médio por serem compostos por aproximadamente 235.000 palavras cada.

Em relação à especificidade do corpus, o autor ressalta que um estudo pode ter um corpus representativo ao focar as obras de um autor somente, ou ainda de uma subárea dentro de um campo geral, como na medicina, por exemplo. Isso ocorre porque “os corpora compilados em pequena escala por pesquisadores individuais acabam sendo mais representativos do que os respectivos sub-corpora dos corpora gerais” (Berber Sardinha, 2000 :348).

São muitas as vantagens de se trabalhar com um corpus eletrônico. Possibilita o gerenciamento automático de milhões de dados em textos contínuos, viabiliza o confronto entre teoria e dados empíricos da língua e

amplia o conhecimento das estruturas lingüísticas por meio da análise dos dados, reproduzindo com considerável fidelidade a realidade lingüística.

Uma das reflexões que nos propomos a realizar neste trabalho é a escolha de uma definição para os termos médicos a serem selecionados para a composição do léxico bilíngüe deste estudo. Observamos em diferentes trabalhos que tomam como base a lingüística de corpus, como o de Castanho (2004), também na área médica, no qual combinações de duas ou mais unidades lexicais são geralmente chamadas de *colocações*. Berber Sardinha (2004:42) explica que a colocação é uma “associação entre itens lexicais, ou entre o léxico e campos semânticos”. Acrescentamos a definição de Manning & Schütze (1999:141, apud Castanho, 2004:2), que consideram a colocação “uma expressão que consiste em duas ou mais palavras que correspondem a alguma forma convencional de se dizer alguma coisa”.<sup>3</sup> As colocações são formadas por uma base e outro elemento que pode ser um substantivo, um verbo ou um adjetivo e que, ao serem combinados, formam uma unidade lexical. No entanto, verificamos que nesses trabalhos, as colocações normalmente seguem um mesmo padrão que é escolhido pelo pesquisador, ou seja, substantivo + adjetivo ou verbo + substantivo etc. Em nossa pesquisa, desejávamos levantar qualquer termo que fosse de relevância e específico da subárea cardiológica, independente da possível combinação que estivesse sendo realizada. Por esse motivo, e por sabermos que na terminologia a combinação de itens lexicais também apresenta diferentes conceituações, decidimos utilizar a proposta de Barros (2004) para termos simples, compostos e complexos. A autora explica que o termo simples é definido como “constituído de um só radical, com ou sem afixos” (ISO 1087, 1990:7) e o termo complexo é “constituído de dois ou mais radicais, aos quais podem-se acrescentar outros elementos” (ISO 1087, 1990:7). Em relação aos termos compostos, Barros considera que:

(...) também são unidades lexicais formadas por dois ou mais radicais. Distinguem-se, no entanto, dos termos complexos pelo alto grau de lexicalização e pelo conjunto de morfemas lexicais e/ou gramaticais que os constitui, em situação de não-autonomia representada graficamente pela utilização do hífen. [...] Cumpre ressaltar que consideramos as unidades lexicais compostas por aglutinação (como fidalgo, embora etc.) e pela justaposição sem hífen de dois ou mais radicais como termos simples (Barros, 2004: 100).

---

3. *A collocation is an expression consisting of two or more words that correspond to some conventional way of saying things.*

No campo da tradução, Tognini-Bonelli (2001) tem realizado investigações voltadas para a noção de unidades funcionalmente completas de significado, extraídas de linhas de concordância<sup>4</sup> a partir do tratamento de corpora computadorizados. A pesquisadora destaca o fato de as palavras não ocorrerem isoladamente, mas sim em suas relações semânticas e funcionais com outras palavras. A respeito das unidades funcionalmente completas de significado, Tognini-Bonelli as define da seguinte maneira:

1. Podem ser identificadas por meio de um processo de formalização de padrões de co-seleção dentro do contexto de uma palavra ou expressão. Envolvem escolhas colocacionais (padrão lexical) e coligacionais (padrão gramatical) e, dessa forma, não podem ser definidas somente em termos lexicais ou em termos gramaticais. Também envolvem uma preferência semântica, percebida por palavras que pertencem ao mesmo campo semântico, e que desempenham uma prosódia semântica específica (associação entre itens lexicais e conotação negativa, positiva ou neutra) nos níveis pragmático e conotativo.
2. São unidades sintagmáticas e por isso inter-relacionam-se com outras palavras. Por meio de um processo de co-seleção, formam uma unidade de multi-palavras que se torna disponível a partir de uma escolha única.
3. Somente quando essas unidades de multi-palavras se apresentam funcionalmente completas é que se tornam disponíveis como equivalentes na tradução ou como unidades comparáveis de significado entre duas línguas.

Outro aspecto abordado refere-se à construção dos corpora para uma pesquisa em tradução. Tognini-Bonelli enfatiza que é importante basear-se em dois corpora comparáveis, um com textos originalmente escritos na língua 1 (L1, no caso do presente trabalho o português) e outro originalmente escritos na língua 2 (L2, nesse caso o inglês) por possibilitarem uma melhor identificação da forma e função das palavras. O primeiro passo nesse tipo de análise consiste na identificação e classificação do padrão

---

4. Por linhas de concordância entende-se “concordâncias ou listagens das ocorrências de um item específico (chamado palavra de busca ou nóculo, que pode ser formado por uma ou mais palavras) acompanhado do texto ao seu redor (o cotexto)” (Berber Sardinha, 2004:105).

formal (padrão colocacional e coligacional) dentro do contexto de uma palavra ou expressão. O segundo passo é o de comparação entre L1 e L2, que permite levantar uma primeira aceção (*prima-face*). Se o analista dispuser, como no caso dessa investigação, de um corpus paralelo, o processo torna-se mais enriquecedor, pois minimizará a dependência da sua intuição. O terceiro passo começa pela função, para observar a forma de realização (padrão coligacional e colocacional) na L2. Esse processo, sugerido por Tognini-Bonelli (2001), tem relação com o processo de decodificação e codificação em outra língua. Com a ajuda dos corpora comparáveis nas duas línguas, o analista terá acesso ao termo como ele é empregado, na L1 e na L2, dentro de um contexto, o que possibilitará uma escolha mais adequada do termo equivalente para sua tradução, baseando-se em evidências reais de uso em ambas as línguas. Seguindo os passos de Tognini-Bonelli (2001) é que nos propusemos à compilação do corpus principal (paralelo) e dos dois corpora de controle (comparáveis), um na L1 e outro na L2.

A construção e o uso de corpora comparáveis, além dos corpora paralelos, também são realizados em outros trabalhos como o de Déjean, Gaussier & Sadat (2002), que utilizam a combinação de diferentes modelos para a extração do léxico bilíngüe a partir de corpora comparáveis. Eles propõem diferentes maneiras de se enriquecer um *thesaurus* com novos termos descobertos a partir de corpora paralelos e sugerem que futuros trabalhos enfatizando a extração da terminologia a partir de corpora comparáveis também sejam realizados.

Nesse sentido, a nossa proposta de extração de termos a partir de corpora paralelos e a verificação da existência dos termos em corpora comparáveis originalmente escritos por autores nativos em ambas as línguas mostra convergência nesse ponto com proposta de Déjean, Gaussier & Sadat (2002).

Outros modelos e abordagens automáticas para a extração de termos complexos e compostos a partir de corpora computadorizados também são propostos por Daille, Gaussier & Langé (1994). No entanto, conforme colocado por Dias & Lopes (2005), algumas destas abordagens não são totalmente promissoras, pois deixam de lado locuções verbais, adjetivas e adverbiais, conjuntivas e preposicionais. Além disso, Dias & Lopes destacam que a filtragem lingüística das unidades polilêxicas candidatas a faze-

rem parte de um glossário deve ser realizada após sua extração estatística ao contrário do que é realizado no trabalho de Daille (1995 *apud* Dias & Lopes, 2005). Eles também reforçam a necessidade de se investigar unidades polilexicais a partir de corpora não tratados, ou seja, não etiquetados, evitando assim a modificação do texto original. Em nossa pesquisa, diferentemente dos trabalhos citados anteriormente, não se mostra necessária a etiquetagem do corpus, o que encontra respaldo na proposta de Dias & Lopes (2005) de se trabalhar com um corpus que não tenha sido modificado.

Acreditamos que a vantagem de se extrair termos de linguagem de especialidade a partir de uma abordagem baseada primeiramente no levantamento de palavras-chave<sup>5</sup>, é que no caso da extração de termos médicos, a lista de palavras-chave excluirá palavras que sejam da língua geral levantando somente palavras estatisticamente relevantes e específicas do corpus de estudo, independente de sua classificação gramatical. Posteriormente, observaremos as linhas de concordância e as listas de agrupamentos lexicais<sup>6</sup> (*clusters*) a partir de cada palavra-chave, assim, extrairemos os termos médicos da subárea e, em seguida, consultaremos um especialista da área observada, neste caso, a área de cardiologia, para que esses termos sejam utilizados no levantamento.

Além dos termos selecionados para compor o léxico, acrescentaremos as linhas de concordância para os termos médicos identificados nos corpora. As linhas de concordância permitem ao consulente observar as palavras que estão associadas aos termos médicos e também as construções frasais mais frequentes do gênero tratado, neste caso, nas línguas portuguesa e inglesa.

## 2. Material e método

Quanto ao *material* utilizado para a construção do corpus principal (paralelo) de cardiologia, coletamos quinze artigos publicados pela revista *Arquivos Brasileiros de Cardiologia*, entre os anos de 2000 e 2004, o qual conta com 53.220 palavras nos textos originais (TOs) e 54.613 nos textos traduzidos (TTs).

---

5. As palavras-chave são palavras “cujas frequências são estatisticamente diferentes no corpus de estudo e no corpus de referência” (Berber Sardinha, 2004:96).

6. As listas de agrupamentos lexicais são “listas de seqüências fixas de palavras recorrentes na concordância (Berber Sardinha, 2004:111).

Para a escolha do *tradutor* dos textos do corpus principal, o critério foi o de selecionarmos um profissional com larga experiência na subárea analisada. Por essa razão, escolhemos artigos de cardiologia traduzidos pela Dra. Stella Maris C. Gandour, que tem formação em medicina e também é mestre em tradução pela PUC do Rio de Janeiro.

Quanto aos *procedimentos* adotados para este estudo, os artigos de cardiologia da revista *Arquivos Brasileiros de Cardiologia* publicados até o ano 2001 já foram escaneados, limpos e salvos em forma de texto sem formatação (txt). Já os artigos de 2002 a 2004 foram diretamente salvos em txt, porque a revista passou a ser disponibilizada na Internet.

Para o *material* dos nossos dois corpora de controle (comparáveis), utilizamos textos na L1 e na L2 retirados de sites que apresentavam textos completos de cardiologia, de mesma natureza dos que havíamos retirado da revista usada no corpus principal (paralelo). Dessa forma, obtivemos um corpus de controle (comparável) escrito originalmente em português, com 233.050 palavras; e outro corpus de controle (comparável) escrito originalmente em inglês com 236.292 palavras. Para uma melhor visualização da proposta de Tognini-Bonelli (2001) aplicada à nossa pesquisa, apresentamos o quadro a seguir:

Quadro: Organização dos corpora segundo Tognini-Bonelli (2001)

<i>Corpus de controle (comparável) na L1 – português</i>	<i>Corpus principal (paralelo) – português? inglês</i>	<i>Corpus de controle (comparável) na L2 – inglês</i>
(artigos escritos originalmente em português na subárea de cardiologia)	(artigos escritos originalmente em português e suas respectivas traduções para o inglês na subárea de cardiologia)	(artigos escritos originalmente em inglês na subárea de cardiologia)

Nos *procedimentos* para a compilação dos dois corpora de controle, os artigos que haviam sido escritos também entre 2000 e 2004, na L1 e na L2, foram coletados da Internet e salvos diretamente em txt.

A seguir, os textos foram processados pelo programa WordSmith Tools versão 3.0 (Scott, 1999) a fim de extrairmos os termos médicos para

a composição do léxico. Para tanto, utilizamos três aplicativos do programa, WordList, KeyWords e Concord. Utilizamos, em primeiro lugar, o aplicativo WordList que nos forneceu uma lista de frequência de palavras, da qual extraímos os termos mais recorrentes dos TOs do corpus principal, bem como dos respectivos TTs. Como passo seguinte, fizemos uma comparação desses termos dos TOs com os seus possíveis correspondentes nos TTs, observando as listas de frequência de ambos sucorpora.

Em seguida, geramos as listas de palavras por frequência, extraídas dos textos originalmente escritos em português (TOPs) e dos textos originalmente escritos em inglês (TOIs) dos dois corpora de controle (comparáveis). Passamos, então, a observar se os termos mais frequentes do corpus principal (paralelo) constavam também como termos mais frequentes nas listas dos dois corpora de controle (comparáveis).

As listas de frequência foram geradas com o intuito de obtermos os termos médicos mais frequentes do corpus principal (paralelo) e dos dois corpora de controle (comparáveis). No entanto, não são todas as palavras de alta frequência que podem ser consideradas termos de uma área especializada. Para confirmarmos se as palavras mais frequentes realmente eram significativas dentro do corpus principal, usamos o aplicativo KeyWords, que nos fornece uma lista de palavras-chave ao comparar as listas de frequência de palavras de um corpus de estudo e de um corpus de língua geral ou de referência. Conforme Berber Sardinha, os componentes principais para uma análise de palavras-chave são dois:

1. Um corpus de estudo [principal], representado por uma lista de frequência de palavras. O corpus de estudo é aquele que se pretende descrever. A ferramenta KeyWords aceita a análise simultânea de mais de um corpus de estudo.
2. Um corpus de referência, também formatado como uma lista de frequência de palavras. Também é conhecido como corpus de controle, e funciona como termo de comparação para análise. A sua função é fornecer uma norma com a qual se fará a comparação das frequências do corpus de estudo. A comparação é feita por meio de uma prova estatística selecionada pelo usuário (qui-quadrado ou *log-likelihood*). As palavras cujas frequências no corpus de estudo forem significativamente maiores segundo o resultado da prova estatística são consideradas chave, e passam a compor uma listagem específica de palavras-chave. (Berber Sardinha, 2004:97).

Como corpus de referência da L1 (português), usamos o corpus da *Folha de São Paulo* – FSP (1997) que contém 39.261.868 palavras. Como

corpus de referência da L2 (inglês), utilizamos o corpus *British National Corpus* – BNC Sampler (corpus do inglês britânico), amostra que apresenta 2.530.849 palavras. Empregamos esses dois corpora de referência por serem de fácil acesso e por servirem para o propósito de extração de palavras-chave em vista de serem mais de 5 (cinco) vezes maior que os corpora de estudo, conforme tamanho sugerido por Berber Sardinha (2004:102) para que um corpus seja usado como referência.

Em primeiro lugar, geramos a lista de frequência do corpus da FSP. Em seguida, retomamos a lista de frequência dos TOs do subcorpus principal para contraste com a lista de frequência de palavras da FSP. O intuito foi gerar uma lista de palavras-chave do subcorpus principal de TOs.

Também utilizamos a lista de frequência de palavras do corpus da FSP (como referência) e a lista de frequência do corpus de controle (comparável) na L1 a fim de obter uma lista de palavras-chave referente ao corpus de controle na L1.

O procedimento foi o mesmo com o corpus BNC Sampler. Primeiro, geramos uma lista de frequência de palavras desse corpus. Em seguida, retomamos a lista de frequência de palavras do subcorpus principal de TTs e, de modo análogo ao realizado anteriormente, comparamos essas duas listas de frequência a fim de obter uma lista de palavras-chave referente ao subcorpus principal de TTs.

Em seguida, usamos a lista de frequência de palavras do BNC Sampler, e a lista de frequência do corpus de controle (comparável) na L2 (inglês) para gerar a lista de palavras-chave referente ao corpus de controle na L2.

Desse modo, obtivemos quatro listas de palavras-chave: uma dos TOs do corpus principal na L1; uma de TOPs do corpus de controle na L1; uma de TTs do corpus principal na L2 e outra de TOIs do corpus de controle na L2.

Depois de observarmos e compararmos as palavras-chave das quatro listas geradas pelo programa, passamos a utilizar o aplicativo Concord que, ao fornecer as linhas de concordância em que as palavras (nódulos) aparecem, permitiu-nos sua observação em seu cotexto, tanto para os TOs e TTs do corpus principal (paralelo) quanto para os TOPs e TOIs dos dois corpora de controle (comparáveis). Por cotexto, entende-se o texto ao redor da palavra de busca, também chamado de nódulo (Berber Sardinha, 2004:105).

#### 4. Discussão e resultados

Para a apresentação, primeiramente, dos dez termos médicos mais frequentes nos TOs e TTs do corpus principal de cardiologia, elaboramos a tabela a seguir, com o número de ocorrências e sua porcentagem.

Tabela 2: Termos mais frequentes no corpus principal (paralelo) de cardiologia

<i>Subcorpus em português (TOs)</i>			<i>Subcorpus em Inglês (TTs)</i>		
WordSmith Tools Word	Freq.	%	WordSmith Tools Word	Freq.	%
1. ARTERIAL	213	0,40	1. VENTRICULAR	218	0,40
2. HIPERTENSÃO	173	0,33	2. HEART	211	0,39
3. MITRAL	132	0,25	3. CORONARY	198	0,36
4. VENTRÍCULO	121	0,23	4. HYPERTENSION	164	0,30
5. VENTRICULAR	118	0,22	5. PULMONARY	157	0,29
6. ATRIAL	114	0,21	6. MITRAL	150	0,27
7. PRESSÃO	91	0,17	7. PRESSURE	148	0,27
8. SISTÓLICA	83	0,16	8. ARTERY	142	0,26
9. ARTÉRIA	82	0,15	9. VALVE	127	0,23
10. ATRESIA	82	0,15	10. ATRIAL	123	0,23

Como o nosso estudo parte da língua portuguesa para a seleção dos termos médicos, escolhemos a primeira palavra mais recorrente do subcorpus principal (paralelo) de TOs, “arterial”, para ilustrar como elaboraremos nosso léxico de cardiologia.

Com o intuito de verificar se havíamos escolhido uma palavra significativa para a composição de nosso léxico, contrastamos a lista de palavras-chave dos TOs do corpus principal (paralelo) com a lista de palavras-chave dos TOPs do corpus de controle (comparável) da L1. Como exemplificação, apresentamos as seis primeiras palavras-chave com as classificações em ordem decrescente conforme a lista de palavras-chave da Tabela 3.

A respeito das informações que constam nessa lista, da esquerda para direita (Tabela 3), Berber Sardinha (2004:103) explica:

- \* Coluna Word: os itens do(s) texto(s).
- \* Coluna Freq: a frequência do item no corpus de estudo.
- \* Coluna <nome do arquivo> %: a porcentagem do item em relação ao total do corpus de estudo.
- \* Coluna Freq: a frequência do item no corpus de referência.
- \* Coluna <nome do arquivo> %: a porcentagem do item em relação ao total do corpus de referência.
- \* Keyness: o resultado da estatística de comparação (*log-likelihood* ou qui-quadrado)
- \* P: o valor da significância estatística atingido pelo resultado da estatística.

Tabela 3: Palavras-chave mais recorrentes do subcorpus principal (paralelo) de cardiologia na L1 em relação ao Corpus da Folha de São Paulo (1997)

WordSmith Tools -- 4/10/2004 09:21:58						
N.WORD	FREQ.	CARDPARP.TXT %	FREQ.	FOLHA97.TXT %	KEYNESS	P
1.ARTERIAL	213	0,40	97	2.429,5		0,000000
3.HIPERTENSÃO	173	0,33	216	1.751,9		0,000000
4.MITRAL	132	0,25	3	1.715,3		0,000000
5.VENTRÍCULO	121	0,23	1	1.587,1		0,000000
6.VENTRICULAR	118	0,22	0	1.559,0		0,000000
7.ATRIAL	114	0,21	0	1.506,2		0,000000

Como podemos notar, as duas primeiras palavras que aparecem na lista da Tabela 3, acima, são “arterial” e “hipertensão”, o que confirma nossa escolha feita anteriormente a partir da lista de frequência dos TOs.

Como passo seguinte, utilizamos a ferramenta Concord, para gerar uma lista de concordâncias do subcorpus principal de TOs com a palavra “arterial”. Apresentamos, a seguir, as vinte primeiras linhas envolvendo esse nóculo:

*Linhas de concordância do nódulo “arterial”, extraídas do subcorpus principal de TOs*

N	Concordance	Set	Tag	Word No.	File	%
1	na figura 1. Em ambos a pressão arterial era elevada tracionando-	s		3.601	rdparp.txt	7
2	mHg; moderada quando pressão arterial diastólica está entre 100-			37.103	rdparp.txt	71
3	kg/m2 apresentavam hipertensão arterial sistêmica. Não restam dú			14.976	rdparp.txt	29
4	27,6% (P= 0,05), tinham pressão arterial controlada. Concl			12.432	rdparp.txt	24
5	os entrevistados tinham pressão arterial controlada no momento d			12.455	rdparp.txt	24
6	9 (17,3%) grave. A hipertensão arterial sistêmica moderada preva			37.363	rdparp.txt	71
7	taram prevalência de hipertensão arterial sistêmica semelhante à d			15.129	rdparp.txt	29
8	são; > 55 anos com hipertensão arterial , de início recente ou des			31.408	rdparp.txt	60
9	Frequência da hipertensão arterial na doença de Chagas cró			36.466	rdparp.txt	69
10	o a resposta da pressão sistólica arterial pulmonar à terapêutica.			46.053	rdparp.txt	88
11	r a evolução da pressão sistólica arterial pulmonar em dois casos.			45.914	rdparp.txt	88
12	lência encontrada de hipertensão arterial sistêmica foi maior nos fu			15.159	rdparp.txt	29
13	orização ambulatorial da pressão arterial (MAPA) para melhor averi			38.334	rdparp.txt	73
14	45. A redução inicial da pressão arterial pulmonar parece mais rela			45.853	rdparp.txt	88
15	s - A elevação aguda da pressão arterial , seja de modo “transitório			2.780	rdparp.txt	6
16	a associação entre a hipertensão arterial sistêmica e característica			12.523	rdparp.txt	24
17	ação dos protocolos hipertensão arterial transitória e hipertensão a			4.631	rdparp.txt	9
18	a aorta. No protocolo hipertensão arterial transitória a tração era ex			3.616	rdparp.txt	7
19	a alta prevalência de hipertensão arterial sistêmica no grupo dos n			14.161	rdparp.txt	27
20	a elevação transitória da pressão arterial . No protocolo hipertensão			4.107	rdparp.txt	8

Ao analisarmos a lista inteira, constatamos que a palavra arterial ocorria, na maioria das vezes, como adjunto adnominal em termos, como: “hipertensão arterial”, “pressão sistólica arterial pulmonar”, “pressão arterial”, “hipertensão arterial sistêmica”, “hipertensão arterial sustentada”.

O próximo passo foi observar se os termos extraídos do subcorpus principal de TOs ocorriam de modo semelhante no subcorpus de controle (comparável) de cardiologia na L1 (português). Para tanto, observamos a lista de palavras-chave extraída da comparação da lista de frequência do corpus de controle de cardiologia na L1 e da lista de frequência do corpus de língua geral da FSP.

A título de ilustração, mostramos, na tabela 4, seis palavras-chave do referido corpus, que aparecem em ordem decrescente conforme sua classificação geral.

Tabela 4: Palavras-chave mais recorrentes do corpus de controle (comparável) de cardiologia na L1 em relação ao Corpus da Folha de São Paulo (1997)

WordSmith Tools -- 4/10/2004 09:25:14						
N.WORD	FREQ.	CORPCO-1.TXT %	FREQ.	FOLHA97.TXT %	KEYNESS	P
9.ARTERIAL	509	0,21	97	4.656,2		0,000000
11.VENTRICULAR	387	0,16	0	3.944,3		0,000000
21.VENTRÍCULO	288	0,12	1	2.921,9		0,000000
22.MITRAL	276	0,11	3	2.779,8		0,000000
26.HIPERTENSÃO	291	0,12	216	2.276,7		0,000000
27.ATRIAL	223	0,09	0	2.272,7		0,000000

Notemos que, nessa lista, constam também as palavras primeiramente selecionadas por nós, a partir do subcorpus principal de TOs.

Ao gerarmos a lista de concordâncias com a palavra “arterial” como nóculo, verificamos que ela ocorria 509 vezes no corpus de controle (comparável) de L1, e 97 vezes no corpus de referência da FSP 97. De modo análogo aos procedimentos realizados anteriormente com o subcorpus principal de TOs, passamos a observar a palavra “arterial” na lista de concordâncias gerada pela ferramenta Concord. Os resultados confirmaram nossas expectativas. Notamos que as palavras que haviam sido usadas pelos autores dos artigos do subcorpus principal de TOs coincidiam com os termos levantados do corpus de controle (comparável) na L1, demonstrando assim que a linguagem médica utilizada em ambos os corpora era semelhante e significativa para uma pesquisa como a que estamos realizando. Obtivemos, por meio do corpus de controle (comparável) da L1, 196 ocorrências do termo “pressão arterial”, 142 ocorrências do termo “hipertensão arterial”, 51 ocorrências de “hipertensão arterial sistêmica”, no entanto, não encontramos nenhuma ocorrência do termo “hipertensão arterial transitória” em nosso corpus de controle (comparável) na L1.

Passamos, então, à observação dos termos traduzidos do subcorpus principal de TTs. Já havíamos realizado o levantamento das palavras mais freqüentes do referido subcorpus, conforme apresentado no início da discussão dos resultados. A seguir, efetuamos o levantamento da lista de palavras-chave do subcorpus principal de TTs. Verificamos que as palavras que haviam aparecido na listas do subcorpus principal de TOs em português também estavam presentes na lista de palavras-chave dos TTs em inglês, como vemos na Tabela 5:

Tabela 5: Palavras-chave mais recorrentes do subcorpus principal (paralelo) de cardiologia na L2 em relação ao Corpus BNC Sampler

WordSmith Tools -- 30/9/2004 18:06:38						
N.WORD	FREQ.	CARDPARI.TXT %	FREQ.	BNC.LST %	KEYNESS	P
2.VENTRICULAR	218	0,40	0	1.608,8		0,000000
4. HYPERTENSION	164	0,30	0	1.210,1		0,000000
7. MITRAL	150	0,27	0	1.106,8		0,000000
12.ATRIAL	123	0,23	0	907,5		0,000000
31.VENTRICLE	74	0,14	0	545,9		0,000000
37.ARTERIAL	68	0,12	0	501,6		0,000000

A seguir, observamos quais os termos que a tradutora do subcorpus principal de TTs de cardiologia havia escolhido para os termos levantados do subcorpus principal de TOs. Para tanto, geramos as linhas de concordância com a palavra *arterial* como nóculo. Apresentamos uma amostra com 20 linhas:

*Linhas de concordância do nóculo “arterial”, extraídas do subcorpus principal de TTs*

N	Concordance	Set	Tag	Word No.	File	%	
1	f the 2 patients, in whom pulmonary	arterial	systolic	pressure could be c	45.152	rdpari.txt	84
2	ividuals with predominantly diastolic	arterial	hypertension (> 120 mmHg).		31.889	rdpari.txt	59
3	d into the native aorta or into a rigid	arterial	system. In this last study, t		5.507	rdpari.txt	10
4	applied. The transient and sustained	arterial	hypertension protocols are s		3.658	rdpari.txt	7
5	renal artery flow. The diagnosis of	arterial	hypertension resulting from r		31.591	rdpari.txt	58
6	stenosis: diagnostic significance of	arterial	hemodinamics. J Cardiovasc		33.957	rdpari.txt	63
7	inutes, as occurred in the sustained	arterial	hypertension protocol, cardia		6.245	rdpari.txt	12
8	follows: dyslipidemias, dietary habit,	arterial	hypertension, smoking, alco		25.560	rdpari.txt	47
9	s been well established in restoring	arterial	patency, in preserving left ve		18.260	rdpari.txt	34
10	in blood pressure. In the sustained	arterial	hypertension protocol, the en		4.203	rdpari.txt	8
11	ardiac adjustments are triggered by	arterial	hypertension and result from		3.292	rdpari.txt	6
12	fferentiating transient and sustained	arterial	hypertension protocols was t		5.140	rdpari.txt	10
13	learance, the evolution of pulmonary	arterial	systolic pressure could be f		46.971	rdpari.txt	87
14	n CM. Casos novos de hipertensão	arterial	, Distrito de Saúde Noroeste		39.411	rdpari.txt	73
15	der AL, Co EE, Sage MD. Coronary	arterial	abnormalities in pulmonary a		23.335	rdpari.txt	43
16	tions, respectively, of the sustained	arterial	hypertension protocol. Res		4.447	rdpari.txt	9
17	ing pressure variation, the “transient	arterial	hypertension protocol (TAH)		2.782	rdpari.txt	5
18	subtraction technique, with selective	arterial	catheterization through femor		32.499	rdpari.txt	60
19	osa DJ. Effect of Chagas disease on	arterial	blood pressure. Am Heart J		39.263	rdpari.txt	73
20	. Doença de Chagas e hipertensão	arterial	primária. Arq Bras Cardiol 1		39.237	rdpari.txt	73

Notamos que dois termos se destacavam: *arterial hypertension protocol*, com 12 ocorrências, e *sustained arterial hypertension*, também com 12 ocorrências. Depois temos *arterial systolic pressure*, e *pulmonary arterial systolic pressure* com 7 ocorrências, e outros termos que não são apresentados nesta amostra. Em alguns termos, verificamos que aparecia a palavra *protocol*, o que nos levou a uma observação mais atenta das linhas de concordância.

Em seguida, passamos à observação do corpus de controle (comparável) na L2, a fim de observamos se os termos traduzidos do subcorpus principal de TTs coincidiam com os termos comumente usados por autores de artigos de cardiologia originalmente escritos em língua inglesa.

Semelhantemente aos passos anteriores, levantamos uma lista de palavras-chave do referido corpus na qual também aparecem cinco palavras que fazem parte das listas anteriores, com exceção da palavra *mitral*, que não se mostra significativa no contraste entre o corpus de controle (comparável) na L2 e o corpus do BNC sampler. Vejamos:

Tabela 6: Palavras-chave mais recorrentes do corpus de controle (comparável) de cardiologia na L2 em relação ao Corpus BNC Sampler

WordSmith Tools -- 30/9/2004 18:12:00						
N.WORD	FREQ.	CCCL.TXT %	FREQ.	BNC.LST %	KEYNESS	P
26. VENTRICULAR	292	0,12	0	1.323,0		0,000000
70. HYPERTENSION	158	0,06	0	715,8		0,000000
169.ARTERIAL	95	0,04	0	430,4		0,000000
277.ATRIAL	62	0,03	0	280,9		0,000000
347.VENTRICLE	49	0,02	0	222,0		0,000000

Ao gerarmos as linhas de concordância, obtivemos os seguintes termos: *arterial disease* 15 vezes; *pulmonary hypertension* com 8 ocorrências; *systolic hypertension* com 7 ocorrências e *arterial pressure* com 6 ocorrências.

Podemos notar que os termos extraídos a partir do corpus de controle (comparável) na L2, de um lado, apresentam algumas diferenças de uso dos termos que aparecem no subcorpus principal de TTs; de outro, apresentam certa correspondência com os termos levantados do corpus de controle (comparável) na L1. Isso deve-se, a nosso ver, a dois fatos. Primeiro, ao fato de os termos da linguagem de especialidade, como a linguagem médica, não serem unívocos, ou seja, existem termos usados de modo adverso por diferentes especialistas da cardiologia para designarem um mesmo referente. No caso dos termos levantados neste trabalho e apresentados a uma especialista da subárea de cardiologia<sup>7</sup>, evidenciou-se o processo em que ocorre omissão de uma palavra no termo denominado elipse lexical. Tal processo segue o princípio de economia lingüística. Este foi o caso do termo “disfunção mitral”, que nos livros da subárea, consta como “disfunção da valva mitral”. A esse respeito, a especialista da área consultada esclarece que a palavra “valva” geralmente é omitida, pois os colegas já sabem que ela está ligada ao termo.

Outro fator que pode ter influenciado essa distinção de termos nos corpora paralelos e comparáveis é o fato de eles serem representativos da amostra coletada para este estudo, isto é, os corpora dessa pesquisa não compreendem todos os artigos já publicados na área, mas, sim uma amostra delimitada de artigos coletados pelas pesquisadoras. Contudo, decidi-

7. Dra. Josélia de Cássia Menin Brandi, especialista em cardiologia há nove anos.

mos incluir tais termos no léxico proposto por nós em vista de apresentarem informações principalmente para o tradutor, que poderá encontrar uma opção correspondente para o uso em seu texto. Lembramos, mais uma vez, que essa pesquisa está voltada, em primeiro lugar, ao levantamento de um conjunto léxico usado por autores de artigos de cardiologia e pelos tradutores da área, com vistas à confecção de um léxico bilíngüe para o uso do tradutor. No entanto, essa pesquisa poderá também servir como um passo preliminar para um estudo mais aprofundado a ser realizado por pesquisadores da lexicologia/lexicografia e terminologia/terminografia.

Apresentamos, no final deste trabalho, uma amostra do léxico em questão, o qual parte das primeiras palavras-chave que aparecem no sub-corpus principal de TOs.

## 5. Conclusão

Esse trabalho, fundamentado numa abordagem interdisciplinar com base nos estudos da tradução baseados em corpus (Baker, 1993, 1995, 1996; Camargo, 2005), nas pesquisas da lingüística de corpus (Berber Sardinha, 2000, 2004; Tognini-Bonelli, 2001) e na terminologia (Andrade, 2001; Barros, 2004) mostra que uma investigação desta natureza é exequível e interessante de se realizar.

As reflexões mostradas são no sentido de apresentar uma proposta de levantamento de termos médicos a partir de um modelo interdisciplinar a fim de elaborar um léxico bilíngüe voltado, principalmente, ao tradutor. Ao avaliarmos outras propostas de levantamentos de termos em diferentes trabalhos científicos, pudemos perceber que a abordagem ora sugerida mostra convergência em outros trabalhos como os de Dias e Lopes (2005) em relação à diversificação de palavras polilexicais e ao de Déjean, Gaussier & Sadat (2002) no que se refere ao uso de um corpus paralelo e de corpora comparáveis para uma pesquisa bilíngüe.

Pudemos verificar que o software WordSmith Tools (Scott, 1999), por meio de suas ferramentas, facilita consideravelmente a análise de corpora eletrônicos, com o processamento de grandes quantidades de informação. As linhas de concordância servem de apoio, esclarecem dúvidas em relação ao termo que se está levantando ao mostrar o cotexto no qual ele se insere.

No entanto, a intuição do analista é parte primordial de cada etapa do processo total de um trabalho como o que estamos propondo. Futuras discussões com profissionais da área de especialidade também podem ser enriquecedoras para esse tipo de pesquisa.

Apresentamos em anexo, uma amostra do léxico que estamos elaborando, o qual parte das primeiras palavras-chave do subcorpus principal de TOs do nosso estudo. Futuramente, pretendemos complementar essa pesquisa utilizando outras subáreas médicas.

Acreditamos, pois, que este estudo, embora inicial e em pequena escala, possa fornecer subsídios a pesquisadores, tradutores e profissionais da área médica no que concerne ao léxico da cardiologia.

Recebido em fevereiro de 2005

Aprovado em abril de 2007

E-mails: [diva@ibilce.unesp.br](mailto:diva@ibilce.unesp.br) [ptppaiva@terra.com.br](mailto:ptppaiva@terra.com.br)  
[xatara@ibilce.unesp.br](mailto:xatara@ibilce.unesp.br)

## REFERÊNCIAS BIBLIOGRÁFICAS

- ANDRADE, Maria Margarida. 2001. Lexicologia, Terminologia: definições, finalidades, conceitos operacionais. In: Ana Maria Pinto de OLIVEIRA, Aparecida Negri ISQUERDO (Org.) *As ciências do léxico: Lexicologia, Lexicografia, Terminologia*. 2. ed. Campo Grande: Editora UFMS.
- BAKER, Mona. 1993. Corpus linguistics and translation studies: implications and applications. In: *Text and technology: in honour of John Sinclair*. Amsterdam/Philadelphia: John Benjamins: 233-250.
- \_\_\_\_\_. 1995. Corpora in translation studies: an overview and some suggestions for future research. *Target*, Amsterdam, John Benjamins, Vol. 7, n. 2:223-243.
- \_\_\_\_\_. 1996. Corpus-based translation studies: the challenges that lie ahead. In: Harold SOMERS (Ed.). *Terminology, LSP and translation studies in language engineering in honour of Juan C. Sager*. Amsterdam/Philadelphia: John Benjamins Publishing Co: 177-186.
- BARROS, Lídia Almeida. 2004. *Curso Básico de Terminologia*. São Paulo: EDUSP.
- BERBER SARDINHA, Tony. 2000. Lingüística de corpus: histórico e problemática. *D.E.L.T.A.*, São Paulo, EDUC, Vol. 16, n. 2: 323 – 367.

- \_\_\_\_\_. 2004. *Linguística de Corpus*. Barueri, SP: Manole.
- BOUTIN-QUESNEL, Rachel. et al. 1985. *Vocabulaire systématique de la terminologie*. Québec, Publications du Québec, 1985 (Cahiers de l'Office de la Langue Française).
- CASTANHO, Rosa. 2004. *Proposta para a elaboração de um glossário de colocações na área médica – subárea hipertensão arterial*. 92f. 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.
- CAMARGO, Diva Cardoso de. 2005. *Padrões de estilo de tradutores: um estudo de semelhanças e diferenças em corpora de traduções literárias, especializadas e juramentadas*. 512 f. Tese (Livre-Docência em Tradução) – Instituto de Biociências, Letras e Ciências Exatas – Unesp, Universidade Estadual Paulista, São José do Rio Preto.
- DAILLE, Béatrice; GAUSSIER, Éric; LANGÉ, Jean-Marc. 1994. Towards automatic extraction of monolingual and bilingual terminology. International Conference On Computational Linguistics Proceedings of the 15th conference on Computational linguistics – Vol 1: 515 – 521. Kyoto, Japan. Disponível em < <http://portal.acm.org> > Acesso em 15 dez. 2006.
- DIAS, Gaël Harry; LOPES, José Gabriel Pereira. 2005. Extração automática de unidades polilexicais para o português. In: *A Língua Portuguesa no Computador*. São Paulo: Mercado de Letras. Disponível em < <http://www.di.ubi.pt/~ddg/publications/BrasilChapterBook2005.pdf> > Acesso em 05 jan. 2007.
- DÉJEAN, Herve; GAUSSIER, Éric; SADAT, Fatia. 2002. An approach based on multilingual thesauri and model combination for bilingual lexicon extraction. International Conference On Computational Linguistics archive. Proceedings of the 19th international conference on Computational linguistics – Vol 1: 1 – 7. Taipei, Taiwan. Publisher Association for Computational Linguistics Morristown, NJ, USA. Disponível em <<http://portal.acm.org>> Acesso em 17 dez. 2006.
- Organisation Internationale de Normalisation. *Terminologie – Vocabulaire*. Genebra: ISO, 1990 (Norme Internationale ISO 1087, 1990).
- SCOTT, Michael. 1999. *WordSmith Tools: version 3.0*. Oxford: Oxford University Press.
- TOGNINI-BONELLI, Elena. 2001. *Corpus linguistics at work*. Amsterdã/Atlanta: John Benjamins.

## ANEXO

## AMOSTRA DO LÉXICO DE CARDIOLOGIA

Corpus de controle na L1 (português)	Subcorpus principal de Tos	Subcorpus principal de TTs	Corpus de controle na L2 (inglês)
<p><i>Pressão arterial</i></p> <p>A monitorização da <i>pressão arterial</i> é criteriosa durante o implante, tendendo a elevar-se durante a expansão dos dispositivos e a cair (...)</p>	<p><i>Pressão arterial</i></p> <p>Resultados - A elevação aguda da <i>pressão arterial</i>, seja de modo “transitório” ou “sustentado”, não alterou a frequência cardíaca e aumentou de maneira equivalente a pré a pós-carga.</p>	<p><i>Arterial pressure</i></p> <p>Results - Acute elevation of <i>arterial pressure</i>, being it “transitory” or “sustained”, did not alter the heart frequency and increased similarly the preload and after load.</p>	<p><i>Arterial pressure</i></p> <p>Angiotensin II nearly doubled the protein synthesis rate in the aorta and small mesenteric arteries, without affecting <i>arterial pressure</i>. The AT1 receptor antagonis</p>
<p><i>Pressão arterial</i></p> <p>A <i>pressão arterial</i> média, controlada pelo anestesista e perfusionista, depende da resistência arterial periférica, devendo ser mantida entre 50-70mmHg.</p>	<p><i>Pressão arterial</i></p> <p>(...) produz uma resposta bifásica na <i>pressão arterial</i> e modula a função endotelial, enquanto não aumenta o peso ou prejudica a homeostase de glicose e insulina.</p>	<p><i>Blood pressure</i></p> <p>(...) produces a biphasic response on <i>blood pressure</i>, and modulates the endothelial function, while it neither increases body weight nor impairs glucose-insulin homeostasis.</p>	<p><i>Blood pressure</i></p> <p>The <i>blood pressure</i> lowering effect of irbesartan was similar in both treatment regimens (Table 1).</p>
<p><i>Hipertensão pulmonar</i></p> <p>(...) <i>hipertensão pulmonar</i>, história de acidente vascular cerebral, doença obstrutiva pulmonar, doença vascular periférica, doença vascular cerebral, características demográficas- idade e sexo;</p>	<p><i>Pressão sistólica arterial pulmonar</i></p> <p>Em dois casos onde foi possível o cálculo da <i>pressão sistólica arterial pulmonar</i> pelo jato da (...)</p>	<p><i>Pulmonary arterial systolic pressure</i></p> <p>Of the 2 patients, in whom <i>pulmonary arterial systolic pressure</i> could be calculated based on the tricuspid regurgitation jet, (...)</p>	<p><i>Pulmonary hypertension</i></p> <p>Exceptions for occlusion include irreversible <i>pulmonary hypertension</i> with right-to-left shunt (PVR 10 units)</p>
<p><i>Valva Mitral</i></p> <p>(...) através da <i>valva mitral</i> em direção ao ventrículo esquerdo</p>	<p><i>Valva mitral</i></p> <p>(...) 33 pacientes com indicação de correção cirúrgica da <i>valva mitral</i> e com fibrilação atrial</p>	<p><i>Mitral Valve</i></p> <p>Thirty-three (67% were women) patients with permanent atrial fibrillation and indication for surgical correction of the <i>mitral valve</i></p>	<p><i>Mitral Valve</i></p> <p>(...) passing through the <i>mitral valve</i> and into the left ventricle(...)</p>