

Colocações especializadas em alemão e português na área de Cardiologia

Leonardo Zilio*

Abstract: This study describes the methodology used in the identification of specialized phraseologies in German and Portuguese with aims at helping translators. To achieve this goal, we adopted Textual Terminology and *Corpus* Linguistics conceptions - among other areas - and offered a methodology of work. This paper focused on specialized collocations that are built by the combination of noun compounds and verbs or deverbatives. Using a *corpus* of Cardiology - in German and Portuguese - we developed a methodology that enables the identification of this type of combination in German as well as the localization of the corresponding equivalent in Portuguese. During our research, we tested statistical methods using calculations such as mutual information and T-score. We also tested contrastive methods based on inter- and intralinguistic contrasts of contexts. The contrastive method presented more satisfactory results. From the *corpus* wordlist, we selected 11 noun compounds as a starting point for our study and then found 97 groups of specialized collocations with their corresponding phraseological equivalents. One example of specialized collocation with equivalent can be seen in the pair *Wiederherstellung von Sinusrhythmus* and *restauração do ritmo sinusal*. In this study, we realized that the use of statistical methods should be carefully analyzed because it may not consider important information, especially when the objective is to help translators.

Keywords: specialized phraseology; language for special purposes; *corpus* linguistics; contrastive studies; cardiology.

Resumo: Este estudo descreve a metodologia utilizada em pesquisa realizada para o reconhecimento de fraseologias especializadas em português e alemão, de forma que se possa auxiliar o tradutor desse par de línguas. Para tal, nos utilizamos do aporte teórico da Terminologia Textual e da Linguística de *Corpus*, entre outras áreas, e propomos uma metodologia de trabalho. O foco deste trabalho foram as colocações especializadas formadas pela associação entre compostos nominais e verbos ou deverbais. A partir de um *corpus* de estudo da área da Cardiologia, em alemão e em português, procuramos desenvolver uma metodologia que permitisse reconhecer esse tipo de associação em alemão e encontrar os respectivos equivalentes fraseológicos em português. No percurso, testamos métodos estatísticos, através do emprego de cálculos como o de informação mútua e o escore T, e métodos contrastivos, nos

*Doutorando do Programa de Pós-graduação em Letras da Universidade Federal do Rio Grande do Sul. Email: lzilio@ig.com.br.

apoiando no contraste inter e intralinguístico de contextos. O método contrastivo apresentou resultados mais satisfatórios. Seleccionamos 11 compostos nominais como ponto de partida para o estudo a partir da lista de palavras do *corpus* e encontramos 97 grupos de colocações especializadas com os respectivos equivalentes fraseológicos. Um exemplo de colocação especializada com equivalente pode ser visto no par *Wiederherstellung von Sinusrhythmus* e *restauração do ritmo sinusal*. Pudemos perceber, com a realização deste estudo, que a aplicação de métodos estatísticos deve ser relativizada, já que pode excluir muita informação importante, principalmente quando se pensa em um tradutor como alvo.

Palavras-chave: fraseologia especializada; linguagem especializada; linguística de *corpus*; estudos contrastivos; Cardiologia.

1. Introdução

O reconhecimento de que não é suficiente para um tradutor saber somente equivalências para determinados nomes, mas também de que é necessário ter informações sobre as combinatórias desses nomes faz com que estudos fraseológicos ganhem em importância. Devido a essa expectativa de fornecimento de unidades mais complexas do que uma lista de substantivos, nos propomos, neste estudo, a oferecer um método de reconhecimento de colocações especializadas.

Nosso objetivo se resume a apresentar dois métodos de reconhecimento fraseológico que foram testados, um estatístico e outro contrastivo. Após essa apresentação, fazemos algumas considerações sobre o tratamento fraseológico, tomando como base uma série de colocações especializadas, de forma a poder auxiliar tradutores especializados que trabalham com textos de Cardiologia¹.

¹ Nosso posicionamento quanto à tradução está de acordo com o que propõe HURTADO ALBIR (2005), que entende tradução como “um processo interpretativo e comunicativo que consiste na reformulação de um texto com os meios de outra língua e que se desenvolve em um contexto social e com uma finalidade determinada”. Não nos aprofundaremos mais na discussão acerca da teoria da tradução, pois acreditamos ser de maior relevância a discussão de outros aspectos teóricos deste estudo.

No percurso para atingir nossos objetivos, usamos principalmente o aporte teórico dos estudos terminológicos e da Linguística de *Corpus*. Os estudos de gêneros textuais (ou tipos textuais) também são referidos aqui, já que proporcionam as bases para um estudo mais aprofundado dos textos especializados de que tratamos.

As áreas de estudos mencionadas anteriormente serão abordadas através de algumas ideias interessantes e polêmicas, levando em conta que não é de nosso interesse abordar com profundidade cada um dos temas, o que nos importa é poder mostrar os resultados de nosso estudo.

A esta introdução, seguem-se sete partes. Começamos apresentando as bases teóricas na Seção 2, abordando o que estamos entendendo por linguística do texto especializado – e também algumas controvérsias associadas. Na Seção 3, continuamos a parte teórica fazendo algumas considerações sobre gêneros textuais. A Seção 4 fecha a parte teórica, discorrendo sobre a Linguística de *Corpus*. Na Seção 5, descrevemos as colocações especializadas sob estudo, e a Seção 6 apresenta o *corpus* utilizado neste estudo, além de fazer uma breve descrição macro e microestrutural dos artigos científicos de Cardiologia. A Seção 7 apresenta e discute os métodos empregados, explicando por que cremos que o método contrastivo foi mais útil para os objetivos deste estudo. Os resultados do método selecionado são apresentados na Seção 8 e a última seção é reservada para algumas considerações e perspectivas do estudo.

2. Terminologia Textual ou Linguística do Texto Especializado

Existem vários autores que já utilizam o nome Terminologia Textual para designar uma teoria que trata de termos em textos (ver CONDAMINES 2005; KRIEGER 2008), já outros usam esse nome para se referir a uma teoria do termo

que reconhece a importância de se observar outros elementos textuais, não só os que o rodeiam, mas os que compõem o texto como um todo (ver LEITCHIK 2002). Neste estudo, tomamos por base uma série de ideias, desenvolvidas principalmente por dois pesquisadores alemães (cf. HOFFMANN 1988a, 1988b; KALVERKÄMPER 1983) que abordam o texto especializado como um todo, reconhecendo os termos somente como uma parte integrante desse objeto multifacetado. Essa perspectiva, não tão nova, porém pouco conhecida no Brasil, é apresentada no País de forma pioneira por FINATTO (2004).

Segundo essa perspectiva, o texto especializado é entendido como:

instrumento e resultado da atividade linguístico-comunicativa desenvolvida em conjunto com uma atividade socioprodutiva especializada. Ele conforma uma unidade estrutural-funcional (totalidade) e constitui-se de uma quantidade finita e ordenada de orações pragmática, semântica e sintaticamente coerentes (textema) ou de unidades com valor oracional que correspondem a enunciados complexos na consciência humana ou a referências complexas na realidade objetiva² (HOFFMANN 1988a).

Como aponta HOFFMANN (1988a), o texto especializado comporta vários níveis de análise, que vão desde o grafema até o textema, passando pelos níveis morfológico, sintático e semântico. A constituição das partes:

se constrói a partir de fonemas, passando por palavras, sintagmas, orações, sequências oracionais, seções/capítulos, texto, comunidades textuais, até gêneros literários (ou gêneros textuais), o que já não pode mais ser observado como uma completude³ (KALVERKÄMPER 1983:153).

² No original: *Instrument bzw. Resultat der im Zusammenhang mit einer spezialisierten gesellschaftlich-produktiven Tätigkeit ausgeübten sprachlich-kommunikativen Tätigkeit. Er bildet eine strukturell-funktionale Einheit (Ganzheit) und besteht aus einer endlichen, geordneten Menge pragmatisch, semantisch und syntaktisch kohärenter Sätze (Texteme) oder satzwertiger Einheiten, die als komplexen Aussagen im Bewußtsein des Menschen und komplexen Sachverhalten in der objektiven Realität entsprechen.* Todas as traduções presentes no texto são de responsabilidade dos autores do artigo.

³ No original: *baut sich auf aus den Phonemen über die Wörter, Syntagmen, Sätze, Satz-Folgen, Abschnitte/Kapitel, Text, Texte-Gemeinschaft, bis in die Gattungen (oder Textsorten) hinein, was als Ganzheit schon kaum mehr überschaubar ist.*

Em nosso estudo, nossa unidade de análise — as colocações especializadas — se encontra principalmente no nível sintático, como veremos mais adiante, mas recorreremos também ao nível semântico para a identificação de equivalentes.

Nessa perspectiva terminológica, a importância do estudo recai sobre os diversos elementos presentes no texto especializado, de forma que os estudos de gêneros textuais também ganham grande espaço, já que permitem a distinção entre as diferentes formas de arranjos de textos. Por consequência, também está envolvido um estudo das linguagens especializadas, pois são elas que conformam o tipo textual e é através delas que se pode percebê-lo.

O termo Terminologia Textual ainda é controverso, tendo em vista que existem vários entendimentos possíveis acerca dele. Já o termo Linguística do Texto Especializado, por outro lado, não apresenta várias possibilidades de entendimento, mas também não é um termo muito conhecido na literatura especializada sobre o assunto no Brasil. Além disso, esse termo parece “esconder” a sua forte ligação com os estudos terminológicos, algo parcialmente desejado, tendo em vista que a “ciência do termo” passa a ser uma “ciência do texto com termos”. Neste estudo, optamos por utilizar a denominação Linguística do Texto Especializado⁴, mesmo sabendo da controvérsia existente, pois cremos que é necessária uma denominação que não esteja associada a outras vertentes existentes dos estudos terminológicos.

⁴ Existem também outras denominações para essa área de estudos — como Linguística de Linguagens Especializadas ou Pesquisa de Linguagens Especializadas —, revelando uma concepção teórica ainda em formação. Para informações mais detalhadas sobre essa questão, pode-se consultar ZILIO (2010).

3. Gêneros ou tipos textuais

Segundo POSSAMAI (2004) — assim como já apontou CIAPUSCIO (2003) — “a expressão *gênero textual* tem seu fundamento atribuído principalmente à conceitualização de *gênero discursivo* na obra de Bakhtin, *Estética da criação verbal* (1997)”. BAKHTIN (1997) propôs a existência dos gêneros do discurso e os definiu como “*tipos relativamente estáveis de enunciados*” (*grifo nosso*). Apesar de gênero discursivo e gênero textual não serem exatamente a mesma coisa, existem vários elementos que apontam para uma grande semelhança entre os dois termos. BALDO (2004) aponta que “embora Bakhtin use a expressão gêneros do discurso e não do texto, nos parece que discurso cobre, para o autor, qualquer tipo de texto oral ou escrito”. Outro autor importante nos estudos de gêneros textuais é SWALES (1990), que, usando apenas a expressão “gênero”⁵, se refere a textos que compartilham um objetivo comunicativo.

Dentro do quadro teórico da Linguística Textual, alguns autores propuseram uma distinção entre classes textuais e tipos textuais. Conforme aponta CIAPUSCIO (1994), classes textuais seriam “classificações empíricas, tais como são realizadas pelos membros de uma comunidade linguística, ou seja, classificações quotidianas (...)”⁶. Além disso, classes textuais podem ser mencionadas por um título comum, como, por exemplo, artigo científico. Já os tipos textuais seriam “uma categoria ligada a uma teoria para a classificação científica de textos”⁷ (CIAPUSCIO 1994). Segundo essa visão, os gêneros textuais ou discursivos seriam entendidos como classes textuais. Já para haver uma tipologia, seria necessário um estudo mais aprofundado dos textos em questão.

⁵ *Genre*, em inglês.

⁶ No original: *clasificaciones empíricas, tal cual son realizadas por los miembros de una comunidad lingüística, es decir, clasificaciones cotidianas (...)*.

⁷ No original: *categoría ligada a una teoría para la clasificación científica de textos*.

Segundo nosso entendimento, a importância dos gêneros textuais está em apontar os textos de nosso *corpus* como pertencentes a um tipo textual em particular. Dessa forma, o artigo científico é entendido como um gênero textual, por ser uma classificação empírica e comumente usada para designar as produções científicas publicadas em periódicos. O artigo científico de Cardiologia, por outro lado, é entendido como um tipo textual, pois recebe um tratamento especial ao ter suas divisões esmiuçadas (como veremos na Seção 6.4).

4. Linguística de *Corpus*

A Linguística de *Corpus* entra neste estudo como norteadora das metodologias testadas e, em última instância, como uma contribuição teórica para o nosso entendimento do que vem a ser linguagem. Existe uma controvérsia quanto ao emprego da Linguística de *Corpus* apenas como uma metodologia de estudo ou como uma abordagem teórica⁸. Nossa abordagem faz bastante uso da parte metodológica da Linguística de *Corpus*, mas nem por isso ignora seus princípios e seu posicionamento teórico.

HOEY (1997; *apud* BERBER SARDINHA 2004) aponta que a “Linguística de *Corpus* não é um ramo da linguística, mas a rota para a linguística”. BIBER *et alii* (1998) também adotam essa perspectiva de abordagem teórica, denominando seu estudo de abordagem baseada em *corpus* e propondo quatro principais características:

- é empírica, analisando os padrões de uso existentes em textos naturais;
- utiliza uma coleção grande e selecionada de textos naturais, conhecida como *corpus*, como base para a análise;
- faz amplo uso de computadores para a análise, utilizando tanto técnicas automáticas como interativas;

⁸ Para mais informações sobre essa questão, pode-se consultar ZILIO (2010).

- depende de técnicas de análise tanto quantitativas quanto qualitativas⁹ (BIBER *et alii* 1998).

Além disso, a Linguística de *Corpus* entende que a linguagem é “um sistema probabilístico de combinatórias” (BERBER SARDINHA 2004). Isso quer dizer que, para cada escolha lexical realizada — ou seja, para cada palavra que se escreve ou fala — existem determinadas probabilidades de que outras palavras (mas não qualquer palavra) a sigam. Esse pressuposto vai contra a ideia de que o ser humano pode escolher livremente as palavras de um texto, explicitando, em vez disso, que as escolhas lexicais futuras são restringidas pelas escolhas já realizadas.

Esse é o pressuposto que nos fez abordar as colocações especializadas a partir de um viés estatístico inicialmente puro, sem observar questões qualitativas, como será descrito na Seção 7.1. Essa abordagem estatística inicial foi posteriormente associada a um viés qualitativo, como descrevemos na Seção 7.2, porém, sempre tivemos em mente o pressuposto das probabilidades de combinatórias.

5. Unidade de análise: colocações especializadas

Para chegarmos a uma delimitação do que seria entendido como a unidade de análise deste estudo, observamos os parâmetros apresentados em vários estudos fraseológicos, tanto de linguagem comum quanto de linguagens especializadas (ver, entre outros, BALLY 1951 [1909]; PICHT 1987, 1990;

⁹ No original: - *it is empirical, analyzing the actual patterns of use in natural texts*; - *it utilizes a large and principled collection of natural texts, known as a 'corpus', as the basis for analysis*; - *it makes extensive use of computers for analysis, using both automatic and interactive techniques*; - *it depends on both quantitative and qualitative analytical techniques*.

HAUSMANN 1989; KJÆR 1990; SINCLAIR 1991; BLAIS 1993; GOUADEC 1994; ROBERTS 1994; MARCO 2000; PAVEL 2003; BEVILACQUA 2004).

Nosso trabalho é focalizado em tradutores, por isso, cremos ser interessante que se usem como ponto de partida do trabalho elementos que constituem uma dificuldade para esse profissional. PICKBRENNER (2005) e LEIPNITZ (2005) mostram em seus estudos que os compostos nominais de língua alemã se apresentam como obstáculos para o tradutor, já que têm envolvidos em sua composição mais do que um substantivo, sem nenhum elemento morfológico ou sintático que permita reconhecer qual a relação estabelecida entre eles. Esse tipo de estrutura, bastante frequente em alemão, geralmente é representada em português por construções sintagmáticas do tipo substantivo + preposição + substantivo ou adjetivo + substantivo; porém, não há como saber de antemão qual a escolha certa, sendo que essas são possibilidades mais frequentes, mas há várias exceções. Exemplos em que se pode aplicar essa conversão são os casos de *Myokardinfarkt* (subst. + subst.), em alemão, e *infarto do miocárdio* (subst. + prep. + subst.), em português, ou *Herzinsuffizienz* (subst. + subst.) e *insuficiência cardíaca* (subst. + adj.).

Para que melhor se possa entender o fenômeno da composição nominal, se pode dizer que “na composição nominal, uma base lexicêmica é acompanhada por um determinante lexicêmico”¹⁰ (WEINRICH 2005). Isso quer dizer que um dos substantivos do composto serve como base, à qual serão acrescentadas as marcas de plural e dos casos morfológicos, e que o(s) outro(s) substantivo(s) serve(m) como determinantes dessa base, acrescentando informações sobre ela.

Devido à dificuldade gerada por esses elementos no momento da tradução, achamos interessante tomá-los como ponto de partida para a verificação de fraseologias. Dessa forma, a partir dos compostos nominais de língua alemã, buscamos elementos relacionados a eles que tivessem origem verbal ou que fossem verbos.

¹⁰ No original: *bei der Komposition des Nomens wird eine lexematische Grundform durch eine lexematische Bestimmungsform determiniert.*

Antes de prosseguirmos com o estudo, são necessárias duas definições: a primeira se refere à fraseologia especializada e a segunda, à nossa unidade de análise. Podemos definir uma fraseologia especializada como um sintagma, contíguo ou não, não necessariamente limitado pela oração ou mesmo pela pontuação, que apresenta um determinado nível de fixação e que ocorre, necessariamente, em textos especializados. Essa fixação não pode ser determinada precisamente, de forma que as fraseologias se estendem por um *continuum* (cf. Bally, 1951) que vai desde os elementos linguísticos fixos até os livres, sendo que estes últimos não são considerados fraseológicos. Já a nossa unidade de análise pode ser definida como uma fraseologia em língua alemã formada por dois elementos fundamentais indispensáveis: um composto nominal e um verbo ou um elemento léxico de origem verbal (substantivo deverbal ou *Partizip I* ou *II* empregado como adjetivo).

Por ter essa característica de ser uma associação verificada, inicialmente, a partir de uma “base” (composto nominal) e um “colocado” (elemento verbal ou deverbal), achamos interessante dar a essa unidade de análise o nome de colocação especializada. Tomamos essa decisão porque esse tipo de unidade formada por “base” e “colocado” já foi chamada anteriormente de “colocação” tanto por HAUSMANN (1989) como por SINCLAIR (1991). Como nosso *corpus* (que será descrito na seção seguinte) é formado por textos especializados, acrescentamos essa característica ao nome da unidade, resultando em *colocações especializadas*.

Algo que deve ficar claro aqui é que estamos usando a denominação *colocações especializadas* apenas para as estruturas extraídas dos textos em alemão, que são a base deste estudo. Às estruturas do português, atribuímos apenas o título de equivalentes fraseológicos (ou, simplesmente, equivalentes), pois não há como saber *a priori* como elas estarão configuradas – ou seja, não há como saber de antemão se elas serão formadas por substantivo + adjetivo, se incluirão algum elemento deverbal, como é a

exigência aqui estabelecida no caso do alemão, ou se apresentarão uma estrutura completamente diferente dessas opções.

6. *Corpus*

O estudo se baseou em um *corpus* bilíngue, que pode ser dividido em dois *subcorpora*: um em alemão e outro em português. O *corpus* é composto por textos de revistas que estavam classificadas, segundo o Portal de Periódicos da Capes, dentro da área de Cardiologia e Doenças Cardiovasculares, e pode ser inserido na categoria de comunicação entre especialistas, na qual “se assume que autor e leitor compartilham da mesma língua e que, quando certas palavras e fraseologias são utilizadas, ambos entendem o significado”¹¹ (PEARSON 1998: 36).

Para a coleta e compilação do *corpus*, levamos em conta as recomendações da Linguística de *Corpus*, observando critérios como o tipo, a representatividade, a extensão, a especificidade e a adequação do *corpus*, e nos baseamos em pesquisas já realizadas na área.

6.1 *Subcorpus* em alemão

O *subcorpus* em alemão era constituído por 493 textos (artigos científicos) extraídos de três periódicos originalmente escritos em alemão e editados na Alemanha, todos vinculados a instituições nacionais e estaduais. Esse *subcorpus* tinha um total de 1.363.286¹² *tokens*¹³. As características individuais dos periódicos podem ser melhor visualizadas na Tabela 1.

¹¹ No original: *it is assumed that author and reader share a common language and that when certain words or phrases are used, each understands what is meant.*

¹² Os dados estatísticos do *corpus* foram observados com a ferramenta WordSmith Tools 4.0 (SCOTT 2004).

Periódico	Período dos textos <i>subcorpus</i>	do	Número de textos	Número de palavras
<i>Herz</i> [Coração]	Mar/2004 Jan/2007	-	170	474.548
<i>Zeitschrift für Herz-, Thorax- und Gefäßchirurgie</i> [Revista de Cirurgia Cardíaca, Torácica e Vascular]	Fev/2002 Jan/2007	-	165	429.466
<i>Zeitschrift für Kardiologie</i> [Revista de Cardiologia]	Jan/2001 Set/2004	-	158	459.272
Total	Jan/2001 Jan/2007	-	493	1.363.286

Tabela 1 - Características dos periódicos alemães.

Procuramos manter um no *corpus*, de forma que cada revista contribui de forma muito parecida para o número de total de *tokens*. Os textos estão todos compreendidos no período que vai desde janeiro de 2001 até janeiro de 2007.

O periódico *Zeitschrift für Kardiologie* apresentava textos redigidos exclusivamente em alemão somente até setembro de 2004. Posteriormente, o seu nome foi modificado para *Clinical Research in Cardiology* e ele passou a aceitar somente trabalhos em língua inglesa.

¹³ Número de palavras presentes nos textos. Cada palavra é definida como um conjunto de caracteres delimitados por espaços em branco e/ou sinais de pontuação.

6.2 Subcorpus em português

O *subcorpus* em português era composto por 490 textos (artigos científicos) extraídos de três periódicos originalmente escritos em português e editados no Brasil, todos vinculados a instituições nacionais e estaduais. Esse *subcorpus* apresentou um total de 1.615.269 *tokens*¹⁴. As características individuais dos periódicos podem ser visualizadas na Tabela 2.

Periódico	Período dos textos <i>subcorpus</i>	Número de textos ¹⁵	Número de palavras
<i>Revista da Sociedade de Cardiologia do Estado de São Paulo</i>	Jan/2004 Out/2006	- 206	697.339
<i>Revista da Sociedade de Cardiologia do Estado do Rio de Janeiro</i>	Jan/2005 Jan/2007	- 112	322.346
<i>Arquivos Brasileiros de Cardiologia</i>	Jan/2005 Fev/2007	- 172	595.584
Total	Jan/2004 Fev/2007	- 490	1.615.269

Tabela 2 - Características dos periódicos brasileiros.

¹⁴ O fato de haver mais de 250.000 *tokens* a mais no *corpus* em português em relação ao *corpus* em alemão não significa, de forma alguma, um problema de balanceamento. Não podemos deixar de considerar que a língua alemã apresenta uma sintaxe diferente da portuguesa, incluindo-se aí sua grande facilidade de criar compostos, diferente do que ocorre no português.

¹⁵ A disparidade perceptível entre o número de textos e, conseqüentemente, o número de *tokens* no *corpus* em português se deu pelo fato de a *Revista da Sociedade de Cardiologia do Estado do Rio de Janeiro* somente apresentar artigos de revisão nos números anteriores a janeiro de 2005. Como os artigos de revisão não estavam previstos na conformação do *corpus*, optamos por deixar que o número de textos dessa revista fosse inferior ao das outras.

Estipulamos aqui o período compreendido entre janeiro de 2004 e fevereiro de 2007 como um dos parâmetros para a seleção dos textos, conforme a disponibilidade *on-line* dos periódicos.

6.3 Arquivos originais dos textos do *corpus*

Os arquivos originais, na versão PDF, dos textos em ambas as línguas foram conservados paralelamente ao *corpus*, que continha arquivos somente em versão TXT. Tomamos essa decisão porque, como o *corpus* não pôde conter figuras, tabelas e resumos em língua estrangeira, acreditamos que seria importante mantermos os arquivos que continham essas informações para casos em que se necessitasse tirar qualquer dúvida ou fazer algum tipo de pesquisa nesses elementos que ficaram de fora do *corpus*.

O uso mais intensivo dos arquivos originais ocorreu nos casos em que se necessitou estabelecer os equivalentes dos compostos, quando o resumo em inglês foi utilizado como *tertium comparationis* provisório (conforme está explicado mais adiante, na Seção 7.2).

6.4 O artigo científico de Cardiologia

Conforme foi anunciado na Seção 3, resumimos aqui os resultados de um pequeno contraste macro e microestrutural realizado entre os artigos científicos de Cardiologia em alemão e português¹⁶. Esse contraste serviu para mostrarmos algumas características mais detalhadas dos textos do *corpus*.

As macroestruturas dos artigos compartilham várias semelhanças, sendo que, em ambas as línguas, percebe-se a sequência clássica dos artigos científicos: introdução (às vezes, com uma subseção para os objetivos),

¹⁶ O estudo completo se encontra em FINATTO *et alii* (2011).

metodologia, resultados, discussão e conclusão. Muitas vezes, a conclusão é apresentada sem haver um subtítulo que a distinga da discussão, mas, nesse caso, ela sempre está nas últimas linhas do texto.

O que se percebe (observando-se também outras áreas, como a Radiologia¹⁷) é que as macroestruturas ficam muito a cargo dos periódicos em que os artigos estão publicados e não tanto a cargo do tipo textual em questão. No entanto, é preciso ressaltar que as decisões dos periódicos também estão influenciadas pelo gênero textual em questão (o artigo científico).

Na parte microestrutural, os textos em alemão e em português distanciam-se bastante. A começar pela extensão dos períodos, temos em média 39,19 palavras por período em português, enquanto os textos em alemão apresentam apenas 23,01 palavras por período (quase a metade). Essa mesma tendência, porém, não se preserva em relação à média de palavras por parágrafo, que é maior em alemão, com 97,62 palavras contra 85,42 palavras nos textos em português.

Quanto à frequência de palavras, observamos uma presença massiva dos substantivos "*Patienten*" e "pacientes", sendo eles os primeiros substantivos da lista de frequência. Acreditamos que a presença, em português, das formas verbais "foi" e "é", e, em alemão, de "ist" e "*werden*"¹⁸, é um bom indício (apesar de não configurar garantia) de que a voz passiva é amplamente utilizada.

Assim, a grande diferença observada ficou por conta dos parágrafos longos e frases curtas do alemão em relação aos parágrafos mais curtos com frases mais longas no português. Em Finatto *et alii* (2011), também é possível observar uma comparação entre os colocados adjetivais nas duas línguas,

¹⁷ Uma breve comparação entre artigos de Cardiologia e de Radiologia pode ser vista em ZILIO (2009).

¹⁸ Ambas as formas verbais do alemão, assim como as apresentadas para o português, podem ser utilizadas como auxiliares para a formação da voz passiva. A tradução de ambas pode ser *grosso modo*, respectivamente, "é" e "ser/tornar-se".

ressaltando diferenças de opção lexical por parte dos autores de cada um dos textos.

Após essa caracterização resumida, mas que aborda diversos aspectos dos textos de Cardiologia presentes no *corpus* sob investigação, passamos agora à descrição dos métodos utilizados para a identificação das colocações especializadas.

7. Em busca do método de trabalho

Para conseguirmos identificar em nosso *corpus* como e quais são as colocações especializadas da Cardiologia em língua alemã e os seus respectivos equivalentes em português, vários passos foram seguidos, os quais serão explicados nesta seção.

Nossa primeira intenção era fazer uma busca por fraseologias especializadas utilizando métodos estatísticos de detecção de colocações. Esse método (apresentado na Seção 7.1) tomou por base dois cálculos estatísticos (informação mútua e score T) com limiares explícitos para decidir quais seriam as colocações especializadas.

O segundo método testado foi o método contrastivo, baseado no contraste de contextos escritos originalmente em alemão e português para averiguação de informações que permitissem apontar uma equivalência contextual. Esse segundo método (apresentado na Seção 7.2) foi o que rendeu melhores resultados e, portanto, será mais bem detalhado.

Para a realização de um trabalho com uma quantidade considerável de palavras, como é o nosso caso, com um *corpus* de cerca de 3 milhões de palavras em 983 textos, é importante que se tenha o apoio de *softwares* desenvolvidos para o tratamento de *corpora* informatizados. Neste estudo, optamos por utilizar o *software WordSmith Tools*, nas versões 3.0 (SCOTT

1999) e 4.0 (SCOTT 2004). Esse *software* possui três ferramentas de tratamento textual, das quais utilizamos a *Wordlist* e a *Concordance*.

De posse dessas ferramentas, o primeiro passo que tomamos foi observar a lista de palavras do *subcorpus* em alemão por ordem de frequência, sem aplicar qualquer filtro. A partir da lista de palavras, escolhemos os compostos nominais que fizeram parte deste estudo. A escolha se deu da seguinte forma: os dois mais frequentes da lista, o menos frequente (considerando um mínimo de 100 ocorrências como limiar inferior¹⁹), e mais oito compostos nominais selecionados aleatoriamente entre esses limiares.

A lista completa dos compostos pode ser vista na Tabela 3 a seguir:

Composto nominal	Frequência
Herzinsuffizienz [insuficiência cardíaca]	914
Risikofaktoren [fatores de risco]	609
Zeitpunkt [momento]	439
Herzerkrankung [doença cardíaca]	378
Kontrollgruppe [grupo controle]	349
Herztransplantation [transplante cardíaco]	326

¹⁹ Esse limiar inferior foi estabelecido de maneira subjetiva. Cremos que, se tomássemos um composto nominal com menos de 100 ocorrências, ele não teria contextos suficientes para averiguar as colocações. Tivemos um problema desse tipo no português, em que a baixa frequência de “angiografia coronariana” (62 ocorrências) não permitiu que fossem identificados muitos equivalentes fraseológicos.

Ejektionsfraktion	221
[fração de ejeção]	
Koronarangiographie	220
[angiografia coronariana]	
Sinusrhythmus	216
[ritmo sinusal]	
Leistungsfähigkeit	213
[capacidade funcional]	
Herzkatheteruntersuchung	101
[cateterismo cardíaco]	

Tabela 3 - Compostos nominais selecionados, com as respectivas frequências, averiguadas na lista de palavras do *subcorpus* alemão.

7.1 Métodos estatísticos

Tínhamos a intenção de trabalhar, neste estudo, com métodos estatísticos de detecção de colocações. Não testamos todos os métodos existentes, algo que não seria possível nem era nosso intento, mas somente aqueles referidos por MICHAEL STUBBS (1995) e BERBER SARDINHA (2004).

Esses autores mencionam o uso de cálculos como informação mútua (I) e escore T (T), além da frequência absoluta para averiguar o que seria realmente uma colocação e não uma mera coocorrência ao acaso. As fórmulas para se encontrar os valores de T e I são as seguintes (cf. STUBBS 1995):

$$T = \frac{\left[\frac{f(n,c)}{N} \right] - \left[\frac{f(n) \times f(c)}{N} \right]}{\left[\frac{f(n,c)}{N} \right]} \quad I = \log_2 \left\{ \frac{[f(n,c) \times N]}{f(n) \times f(c)} \right\}$$

onde $f(n,c)$ é a frequência de coocorrência entre a palavra de busca (o núcleo) e colocado; $f(n)$ é a frequência do núcleo; $f(c)$ é a frequência do colocado; e N é o tamanho do *corpus*²⁰.

Para melhor exemplificarmos, podemos tomar dois exemplos de candidatos a colocações testados: *Herzinsuffizienz* [insuficiência cardíaca] (1.067) + *bedingen* [causar] (657) e *Herzkatheteruntersuchung* [cateterismo cardíaco] (152) + *ergeben* [resultar] (259)²¹. Em nosso estudo, os núcleos são sempre os compostos nominais. Nestes casos, o valor de $f(n,c)$ para *Herzinsuffizienz* + *bedingen* é 6, ou seja, essas duas palavras coocorriam seis vezes, já no caso de *Herzkatheteruntersuchung* + *ergeben*, o $f(n,c)$ é igual a 4. Os valores de $f(n)$ são, respectivamente, 1.067 e 152 para *Herzinsuffizienz* e *Herzkatheteruntersuchung*, e os valores de $f(c)$ são, respectivamente, 657 e 259 para *bedingen* e *ergeben*. O valor de N é fixo, por ser o tamanho do *corpus*, que é de 1.363.286 palavras.

Ambos os autores apontam que, para obter informações satisfatórias, o valor de I deveria ser maior ou igual a 3, e os valores de T e de $f(n,c)$, maiores ou iguais a 2.

BENEDUZI (2008) também testou esses dois métodos estatísticos, entre outros, para avaliar as colocações. Os parâmetros estabelecidos pela autora para o que viria a ser reconhecido como colocação era diferente dos que aqui

²⁰ Para realizar esses cálculos, o tamanho do *corpus* utilizado foi o tamanho de cada um dos *subcorpora* e não o tamanho total do *corpus*, já que estávamos trabalhando com um idioma de cada vez.

²¹ Os valores entre parênteses referem-se às frequências. Ressalta-se aqui que a diferença nas frequências de *Herzinsuffizienz* e *Herzkatheteruntersuchung* em relação àquelas apontadas na Tabela 3 se deve ao fato de que a busca foi realizada utilizando-se não só as formas no singular, mas também no plural e com possíveis declinações, enquanto a lista da Tabela 3 apresenta a frequência somente daquela forma específica.

abordamos, porém, as conclusões sobre os cálculos estatísticos foram parecidas.

Para os testes estatísticos, utilizamos os dois compostos já mencionados (*Herzinsuffizienz* e *Herzkatheteruntersuchung*). Caso os resultados fossem satisfatórios, aplicaríamos os cálculos para os outros compostos. Escolhemos justamente os dois extremos de nossa lista de compostos, o mais frequente e o menos frequente.

Os resultados da aplicação dos testes foram bastante frustrantes. De forma resumida, para ambos os compostos, qualquer candidato a colocação especializada com $f(n,c)$ menor do que 5 era refutado automaticamente pelo escore T, que apresentava valores inferiores a 2; em oposição a isso, as coocorrências de frequência 5 eram aprovadas pelo escore T^{22} e pela informação mútua. Dessa forma, os cálculos tornavam-se praticamente inúteis, já que se poderia dizer quais iriam ser aprovados ou não somente pelo valor de $f(n,c)$.

Além de os resultados se mostrarem bastante pobres em relação ao que se esperava, esse tipo de cálculo despreza totalmente a semântica e as palavras gramaticais que também estão presentes em colocações especializadas.

Tendo observado essas limitações nos testes estatísticos, nem mesmo tentamos aplicá-los ao corpus em português. Em vez disso, preferimos partir diretamente para as observações contrastivas, as quais levaram em conta, além de frequência, elementos semânticos e sintáticos, como pode ser visto na próxima seção.

Ter abandonado os métodos estatísticos também não se deu sem um ônus, já que a observação dos contextos foi muito mais extensa, levando-se em conta que muitos dados deixaram de ser excluídos pelos cálculos e

²² Havia possibilidade de refutação de uma colocação com $f(n,c)$ igual a 5, mas, para isso, os colocados precisavam ter uma frequência muito alta. Os colocados de *Herzinsuffizienz* que coocorriam cinco vezes precisavam de frequências superiores a 650 [$f(n,c) > 650$] e os de *Herzkatheteruntersuchung*, superiores a 4.500 [$f(n,c) > 4.500$].

acabaram entrando para a análise. É claro que poderíamos simplesmente ter estipulado um critério de frequência mínimo de coocorrência, porém, tínhamos a crença de que mesmo as baixas frequências podem esconder informações importantes, principalmente pelo fato de que estamos levando em conta critérios semânticos e contrastivos.

7.2 Método contrastivo

O método contrastivo aqui empregado não se baseia em nenhum autor específico, já que, pelo que pudemos observar na bibliografia a que tivemos acesso, não havia nenhum método que se encaixasse ou pudesse ser adaptado para o estudo aqui proposto. Dessa forma, nesta seção do estudo, estamos propondo uma metodologia de contraste.

Tendo deixado de lado os testes estatísticos, precisamos estipular alguns critérios de seleção. O primeiro critério foi, então, que houvesse, no mínimo, duas ocorrências da associação – em termos estatísticos: $f(n,c) \geq 2$. Para não utilizarmos somente a frequência, acrescentamos um critério de distribuição, ou seja, a associação deveria ocorrer, no mínimo, em dois artigos científicos diferentes. Para completar, deveria ter um equivalente fraseológico²³ averiguado em português. Os três critérios foram, portanto, os seguintes:

1. $f(n,c) \geq 2$.
2. Estar presente em mais de um artigo.
3. Ter um equivalente em português.

Através da ferramenta *Concord*, obtivemos as linhas de concordâncias de cada um dos compostos nominais selecionados. Para gerarmos as linhas de

²³ “Equivalente” aqui, assim como no caso do equivalente fraseológico, é entendido como aquela palavra ou sintagma que, em virtude do contexto e do tipo textual em que está empregado, representa uma possibilidade de tradução que não causaria estranheza ao tipo de leitor para o qual se destina o texto traduzido.

concordâncias, não fizemos distinção entre as formas singular e plural do composto, sendo que ambas foram contempladas no estudo, incluindo-se eventuais declinações. Existem estudos apontando para o fato de que o emprego do singular e do plural se dá de forma diferente em muitos casos (ver STUBBS 2001; SINCLAIR 1991); por isso, preferimos observar os dois justamente para podermos perceber e indicar tal comportamento diferenciado, caso haja algum indício de sua ocorrência. Ainda no que diz respeito às palavras de busca, elas foram utilizadas, em alemão, sempre propiciando que se encontrassem os compostos nominais buscados dentro de outros compostos nominais. Assim, ao buscarmos o composto *Herzinsuffizienz* no *corpus*, utilizamos a expressão **herzinsuffizienz**, que nos permite averiguar os elementos à direita e à esquerda junto à palavra, de forma que também foram levados em conta, por exemplo, *Linksherzinsuffizienz* [insuficiência cardíaca esquerda] e *Herzinsuffizienzmodifikation* [modificação da insuficiência cardíaca].

Cada contexto de cada um dos compostos foi verificado individualmente para se observarem os elementos sintáticos e separado de acordo com os colocados potenciais encontrados em torno dos compostos nominais. Por exemplo: todos os contextos que apresentavam *Behandlung der Herzinsuffizienz* [tratamento da insuficiência cardíaca] e *behandelte Herzinsuffizienz* [insuficiência cardíaca tratada] foram salvos em um mesmo arquivo de texto. Dessa forma, ficou mais fácil a comparação entre os contextos para que se pudessem detectar os significados de cada uma das expressões em alemão. A decisão de juntar os contextos com base nos lemas (por exemplo, o colocado *behandeln* [tratar] abriga também *behandelt* [tratado] e *Behandlung* [tratamento]) encontra fundamento em STUBBS (2001), que explica a diferença feita na Linguística de *Corpus* entre lema e formas de palavras (*lemmas and word-forms*, cf. STUBBS 2001). Entende-se que um lema pode abrigar uma série de formas de palavras, sejam elas formas verbais flexionadas ou deverbais. O critério de lematização é complexo e,

naturalmente, “envolve muitas decisões”²⁴ (STUBBS 2001). Neste estudo, decidimo-nos por tomar como lema a forma verbal, que serviu como um agregador das demais formas dela derivadas. Nos casos em que o verbo não ocorria junto ao composto, utilizamos como lema as formas substantivas e adjetivas, respectivamente.

O motivo de utilizarmos a lematização foi a crença de que as diferentes formas derivadas de uma mesma raiz, em muitos casos, somente representam uma escolha diferente, seja ela feita pelo autor ou forçada pelo tipo textual, para se referir a um fato ou uma situação semelhantes. Há de se ter em mente, porém, que, em alguns casos, uma forma de palavra apresentou significado e equivalente diferentes de outra, mesmo que ambas pertencessem a um mesmo lema. Exemplo disso foi *Fortschreiten der Herzinsuffizienz* e *fortgeschrittene Herzinsuffizienz*, em que ambas as formas de palavras dos colocados foram alocadas sob o lema *fortschreiten*, porém, ambas apresentam tanto significados quanto equivalentes distintos: *progressão da insuficiência cardíaca* (entre outros) e *insuficiência cardíaca avançada*, respectivamente. Nesses casos, as formas de palavras foram separadas de acordo com os significados que apresentavam, sem serem mostradas sob o mesmo lema.

Depois de termos juntado os contextos em alemão de acordo com a forma lematizada dos colocados, partimos para a observação dos contextos em português. Para isso, foi necessário, primeiramente, estabelecer os equivalentes em português para cada um dos compostos nominais investigados.

Para confirmarmos os equivalentes sobre os quais já tínhamos conhecimento ou, pelo menos, alguma informação prévia, simplesmente usamos a conhecida técnica de tentativa e erro. Um possível equivalente era inserido na ferramenta *Concord* do *software WordSmith Tools* e se averiguava sua pertinência a partir dos contextos retornados pela ferramenta.

²⁴ No original: *involves many decisions*.

Nos casos mais complicados, utilizamos os resumos em língua estrangeira dos artigos em formato PDF, já que tanto os artigos em alemão quanto os artigos em português apresentavam resumo em inglês, de forma que, em alguns casos, o inglês serviu como *tertium comparationis* provisório. Os resumos em inglês, apesar de não entrarem para o *corpus*, permaneciam nos arquivos originais dos artigos de Cardiologia, de forma que puderam ser acessados através do *software* Acrobat Reader.

É claro que, como não podia deixar de ser, os equivalentes de cada um dos compostos nominais foram verificados em contextos e a prova final de equivalência sempre foi tirada a partir das colocações e das equivalências fraseológicas.

Tendo estabelecido os possíveis equivalentes para os compostos nominais em alemão, cada uma das linhas de concordância do *subcorpus* em português foi perscrutada em busca de candidatos a equivalentes. Os contextos dos candidatos a equivalentes foram, por sua vez, separados em arquivos diferentes, de acordo com os lemas dos colocados que apresentavam.

Após todos os contextos que encontramos em alemão e em português terem sido separados de acordo com os lemas dos colocados neles presentes, partimos para uma avaliação minuciosa desses contextos. Os contextos nas duas línguas foram pareados, buscando-se elementos que apontassem para descrições textuais de situações ou fatos semelhantes nas duas línguas, de forma que se pudesse dizer que uma forma de se expressar em português fosse equivalente à forma existente em alemão.

Como não somos cardiologistas, precisamos, nessa etapa do trabalho, confiar em nossa competência como linguistas conhecedores de mais de um idioma e especialistas no campo da linguagem para julgar os contextos. Não defendemos, porém, que um linguista deva sempre fazer esse tipo de trabalho sozinho. O apoio de um cardiologista, no caso de nosso estudo, seria muito bem-vindo.

Como essa parte do estudo é bastante delicada, já que envolve um julgamento subjetivo²⁵ acerca das questões de equivalência entre sintagmas, buscamos tomar muito cuidado na avaliação dos contextos, procurando sempre elementos que nos pudessem dar indicativos de sua semelhança. Esses elementos eram, em geral, descrições acerca de procedimentos, doenças, sintomas etc. Um exemplo disso pode ser visto a seguir:

an Herzinsuffizienz leiden vs. insuficiência cardíaca acomete

Weltweit leiden mehr als 22 Mio. Menschen an einer Herzinsuffizienz. [No mundo todo, mais de 22 milhões de pessoas apresentam insuficiência cardíaca.]	A insuficiência cardíaca é uma síndrome clínica complexa, acometendo aproximadamente 4,7 milhões de pessoas nos Estados Unidos.
---	---

Houve contextos que realmente não puderam ser identificados como pertencentes a um determinado lema, seja no caso de bases, colocados ou equivalentes, de forma que nem todos os contextos foram aproveitados no estudo. Geralmente, esses contextos excluídos estavam ou incompletos (por configurarem títulos, estando fora de um contexto mais explicativo) ou quebrados (por erros na compilação do *corpus*, algo difícil de evitar em sua totalidade).

²⁵ Por ser um trabalho que envolve avaliação dos significados de expressões em duas línguas, a subjetividade foi um elemento presente ao longo de sua execução, principalmente pelo fato de que semântica não se deixa avaliar com facilidade de maneira objetiva.

8. Resultados do método contrastivo

Após termos escolhido o método contrastivo como sendo o mais apropriado para o estudo, passamos à observação dos dados, chegando a um total de 97 grupos²⁶ de colocações especializadas em alemão com equivalências em português.

Percebemos que mesmo colocações de frequência baixa têm sua importância para o tradutor quando encontram equivalências em português, como foi o caso de *als (eine) Ejektionsfraktion definiert werden ser definida como (uma) fração de ejeção*, em que tanto a colocação especializada quanto o equivalente fraseológico apresentaram duas ocorrências.

Existem casos em que há uma preferência clara por um tipo de colocação especializada, como, por exemplo, no grupo a seguir, que apresenta colocações especializadas em alemão, seguidas por equivalentes fraseológicos em português:

Durchführung der (einer) Koronarangiographie (4)²⁷
 Koronarangiographie durchführen (1)
 Koronarangiographie wurde durchgeführt (37)
 durchgeführte Koronarangiographie (3)
 Koronarangiographie erfolgte (7)

e

realização de angiografia coronariana (4)
 realizar angiografia coronariana (4)
 angiografia coronariana foi realizada (6)
 angiografia coronariana realizada (1)

Neste grupo, *Koronarangiographie wurde durchgeführt* tem frequência muito mais alta do que seus sinônimos, podendo ser considerada a forma

²⁶ Cada grupo representa um conjunto de colocações especializadas, com suas sinonímias, separadas conforme seu significado.

²⁷ Valores entre parênteses representam a frequência no *corpus*.

preferencial. Já no português, não se verifica uma preferência acentuada entre as opções (exceto pela última opção, que teve apenas uma ocorrência e foi considerada porque seu contexto se mostrou similar aos demais contextos do mesmo lema – “realizar”).

Devido ao grande número de colocações especializadas averiguado, pensávamos em utilizar critérios de corte de colocações especializadas, como é proposto por autores que trabalham com o critério de comutabilidade (ver PICT 1990; PAVEL 2003). Porém, observando os grupos de colocações especializadas e suas idiosincrasias, percebemos o quão importante é a observação dos critérios de risco e de vantagem de manipulação propostos por GOUADEC (1994), que, mesmo não sendo precisos e objetivos, devem ser levados em consideração quando tratamos de um conhecimento especializado a ser disponibilizado para tradutores. Assim, decidimos não adotar critérios de corte mais elevado e apresentar todos os casos confirmados em contexto que apresentassem mais de duas ocorrências.

Por esse mesmo motivo, mantivemos, em português, os casos de sinonímia entre equivalentes fraseológicos, para que os consulentes de nosso estudo pudessem ter opções de tradução para casos como o seguinte:

Bei allen Patienten wurde ein Lävokardiogramm mit selektiver Koronarangiographie nach der Judkins-Technik **durchgeführt**. Die Koronarangiographie **erfolgte** in standardisierten Projektionen (*grifo nosso*).

Neste caso, percebe-se claramente que o autor não utilizou a colocação especializada preferencial *Koronarangiographie wurde durchgeführt*, pois essa estrutura verbal já havia sido utilizada na sentença anterior, restando optar por outro verbo: *erfolgen*. No caso dos equivalentes, as diferentes opções servem para suprir essas opções, permitindo uma flexibilidade do vocabulário empregado na tradução.

Dentre os grupos de colocações e equivalentes, percebemos que alguns deles apresentaram diferenciação devido à preposição existente entre a base e o colocado. Um desses casos foi o de *Fortschreiten zur Herzinsuffizienz* [progressão para insuficiência cardíaca] e *Fortschreiten der Herzinsuffizienz* [progressão da insuficiência cardíaca], em que a primeira opção indica a evolução de certo estado para o estado de insuficiência cardíaca, enquanto a segunda indica a evolução da própria insuficiência cardíaca, seja para outro estado, como o do infarto do miocárdio, seja para um estágio mais avançado da doença. O fato de haver relações gramaticais diferentes entre base e colocado fez com que o significado fosse diferente, pois as preposições também carregam um significado em muitos casos.

9. Considerações finais e perspectivas

A partir do método contrastivo de descrição de colocações especializadas, conseguimos apontar fenômenos da linguagem da Cardiologia que permitem auxiliar o tradutor em sua árdua tarefa.

Por enquanto, métodos estatísticos devem ser relativizados na observação de colocações especializadas quando o trabalho é feito em mais de um idioma, já que há outros fatores importantes além da frequência que contribuem para o reconhecimento de uma colocação especializada. Um desses fatores é a existência de vários contextos em mais de um idioma, o que permite um contraste inter e intralinguístico. Enquanto os métodos estatísticos de avaliação de colocações não forem revistos, parece que seu emprego permanece bastante restrito. Ainda assim, faz-se a ressalva de que é possível que um *corpus* maior ou um refinamento nos cálculos forneçam melhores perspectivas.

Através dos grupos de colocações especializadas e equivalentes fraseológicos verificados, percebeu-se que mesmo frequências relativamente

baixas, como no caso de *als (eine) Ejektionsfraktion definiert werden*, com somente duas ocorrências, são importantes para o estudo, já que podem mostrar sua pertinência no contraste interlinguístico.

O fato de termos encontrado casos em que preposições (palavras consideradas gramaticais) fizeram diferença no significado e, conseqüentemente, na equivalência reforça nossa opção pelo método de contraste entre contextos e não por uma avaliação estatística. Isso porque esta última, ao observar somente a base e o colocado (palavras ditas lexicais), teria deixado passar a influência exercida pelas preposições.

Por fim, cabe ressaltar que os dados levantados neste estudo estão sendo organizados em um glossário *on-line* que poderá em breve ser acessado no *site* CardioTrad (<http://www6.ufrgs.br/termisul/cardiotrad/>)²⁸. Ao todo, foram 97 grupos de colocações especializadas e equivalentes fraseológicos, sendo 151 colocações especializadas em alemão e 142 equivalentes fraseológicos em português.

10. Referências bibliográficas

- BAKHTIN, M. *Estética da criação verbal*. 2ª ed. São Paulo: Martins Fontes, 1997 [em russo: 1979]. Traduzido do francês por Maria Ermantina Galvão G. Pereira.
- BALLY, C. *Traité de stylistique française*. 3ª ed. Paris: Klincksieck, 1951. [1909].
- BALDO, A. Gêneros discursivos ou tipologias textuais? In: *Revista Virtual de Estudos da Linguagem - ReVEL*, ano 2, n. 2, mar 2004. Disponível em: <www.revel.inf.br>.
- BENEDEZI, R. *Colocações substantivo + adjetivo: propostas para sua identificação e tratamento lexicográfico em dicionários ativos português-espanhol*. Porto Alegre: UFRGS, 2008. Dissertação de Mestrado.

²⁸ O *site* está passando atualmente por reformas visuais e estruturais, de modo que é possível que o *link* seja modificado. Porém, certamente haverá um *link* na página do Termisul (<http://www6.ufrgs.br/termisul/>).

- BERBER SARDINHA, T. *Linguística de Corpus*. Barueri: Manole, 2004.
- BEVILACQUA, C.R. *Unidades Fraseológicas Especializadas Eventivas: descripción y reglas de formación en el ámbito de la energía solar*. Barcelona: IULA/UPF, 2004. Tese de doutorado.
- BIBER, D.; CONRAD, S.; REPPEN, R. *Corpus Linguistics: Investigating language structure and use*. Cambridge: Cambridge University Press, 1998.
- BLAIS, E. Le phraséologisme. Une hypothèse de travail. In: *Terminologie Nouvelles*, n. 10. Hull: Rint, 1993, p. 50-56.
- CIAPUSCIO, G. E. *Tipos textuales*. Buenos Aires: Universidade de Buenos Aires, 1994.
- CONDAMINES, A. Linguistique de corpus et terminologie. In: *Langages*, 157, 2005, p. 36-47. Disponível em: <http://w3.univ-tlse2.fr:8880/erss/index.jsp?perso=acondami&subURL=Langagesdef.pdf>.
- FINATTO, M. J. B. Termos, textos e textos com termos: novos enfoques dos estudos terminológicos de perspectiva lingüística. In: ISQUERDO, A. N.; KRIEGER, M. G. (Orgs.). *As ciências do léxico: lexicologia, lexicografia, terminologia*, volume II. Campo Grande: Ed. UFMS, 2004.
- _____ et al. Artigos de Cardiologia em português e alemão: contribuições da pesquisa em corpus para o ensino de leitura instrumental. In: VIANA, V.; TAGNIN, S. E. O. (Orgs.) *Corpora no ensino de línguas estrangeiras*. São Paulo: HUB Editorial, 2011, p. 198-227.
- GOUADEC, D. Nature et traitement des entités phraséologiques. In: *Terminologie et phraséologie. Acteurs et aménageurs. Actes du deuxième Université d'Automne en Terminologie*. Paris: La Maison du Dictionnaire, 1994, p. 164-193.
- HAUSMANN, F.J. Le Dictionnaire de Collocations. In: HAUSMANN, F. J.; REICHMANN, O.; WIEGAND, H. E. U. A. (Orgs.). *Wörterbücher, Dictionaries, Dictionnaires. Ein internationales Handbuch zur Lexikographie*. Berlin, New York: Walter de Gruyter, 1989, p. 1010-1018.
- HOFFMANN, L. *Vom Fachwort zum Fachtext: Beiträge zur Angewandten Linguistik*. Tübingen: Gunter Narr Verlag, 1988a.
- _____. Grundbegriffe der Fachsprachenlinguistik. In: *Germanistisches Jahrbuch für Nordeuropa*. 7. Folge. Deutsche Fachsprachen in Forschung und Lehre. Helsinki, Estocolmo, 1988b, p. 9-16.
- HURTADO ALBIR, A. Aquisição da competência tradutória: aspectos teóricos e didáticos. In: PAGANO, A.; MAGALHÃES, C.; ALVES, F. (Orgs.). *Competência em tradução: cognição e discurso*. Belo Horizonte: Ed. UFMG, 2005, p. 19-57.

- KALVERKÄMPER, H. Textuelle Fachsprachen-Linguistik als Aufgabe. In: *Zeitschrift für Literaturwissenschaft und Linguistik*, v. 51/52, n. 13, 1983, p. 124-166.
- KJÆR, A. L. Phraseology research - State-of-the-art: Methods of describing word combinations in language for specific purposes. In: DRASKAU, Jennifer (Org.). *Journal of the International Institute for Terminology Research - ITTF - Terminology Science and Research*. Vol. 1, n. 1-2, 1990.
- KRIEGER, M. G. Porque Lexicografia e Terminologia: relações textuais? In: FINGER, I.; COLLISCHONN, G. (Orgs.) *Anais do 8º Encontro do Círculo de Estudos Lingüísticos do Sul - CELSUL*, Porto Alegre, de 29 a 31 de outubro de 2008. Pelotas: EDUCAT, 2008.
- LEIPNITZ, L. *Compostos nominais em língua alemã em medicina em tradução para o português*. Porto Alegre: UFRGS. Mestrado em Teorias do Texto e do Discurso - Universidade Federal do Rio Grande do Sul. Orientador: Maria José Bocorny Finatto, 2005.
- LEITCHIK, V. M. Elements of Terminological Text Theory. In: SHELOV, S. D.; LEITCHIK, V. M. (eds.) *Russian Terminology Science (1992-2002)*. Viena: Termnet Publisher. Editado com a colaboração de H. Picht e C. Galinski, 2002.
- MARCO, M. J. L. Collocational frameworks in medical research papers: a genre-based study. In: *English for Specific Purposes*, v. 19, n. 1, 2000, p. 63-86.
- PAVEL, S. A fraseologia na língua de especialidade. Metodologia de registro nos vocabulários terminológicos. In: FAULSTICH, E.; ABREU, S. (Orgs.) *Lingüística aplicada à Terminologia e à Lexicologia*. Porto Alegre: UFRGS/NEC, 2003, p. 99-131. Artigo publicado originalmente em 1993, traduzido do francês por Germana H.P. de Sousa.
- PEARSON, J. *Terms in context*. Amsterdã, Filadélfia: John Benjamins Publishing Company, 1998.
- PICHT, H. Fachsprachliche Phraseologie - die terminologische Funktion von Verben. In: CZAP, H.; GALINSKY, C. (Orgs.). *Terminology and Knowledge Engineering*. Proceedings. International Congress on Terminology and Knowledge Engineering, 29 de setembro - 1º de outubro de 1987. Frankfurt a.M.: INDEKS Verlag, 1987, p. 21-34.
- _____. LSP PPhraseology from the terminological point of view. In: DRASKAU, Jennifer (Org.). *Journal of the International Institute for Terminology Research - ITTF - Terminology Science and Research*. vol. 1, n. 1-2, 1990.
- PICKBRENNER, M. B. *Termos compostos em língua alemã: uma contribuição para o ensino de leitura instrumental em direito*. Dissertação (Mestrado em

Teorias do Texto e do Discurso) - Universidade Federal do Rio Grande do Sul. Orientador: Maria José Bocorny Finatto, 2006.

- POSSAMAI, V.; LEIPNITZ, L. Os estudos de gênero e a tradução: uma relação proveitosa demonstrada por meio da abordagem da tradução de artigos científicos. In: *Anais do 4º Simpósio Internacional de Estudos de Gêneros Textuais*, Universidade do Sul de Santa Catarina, Tubarão, SC, Brasil - de 15 a 18 de agosto, 2007, p. 2016-2027. Disponível em: <<http://www3.unisul.br/paginas/ensino/pos/linguagem/cd/index1.htm>>.
- ROBERTS, R. P. Identifying the Phraseology of Languages for Special Purposes (LSPs). In: *Alfa. Actes de langue française et linguistique*, v. 7/8, 1994, p. 61-74.
- SCOTT, M. *WordSmith Tools version 3*. Oxford: Oxford University Press, 1999.
- _____. *WordSmith Tools version 4*. Oxford: Oxford University Press, 2004.
- SINCLAIR, J. *Corpus, Concordance, Collocation*. Oxford: Oxford University Press, 1991.
- STUBBS, M. Collocations and semantic profiles: On the cause of the trouble with quantitative studies. In: *Functions of language*, vol. 2, n. 1. Amsterdã: John Benjamins, 1995.
- _____. *Words and Phrases: Corpus studies of lexical semantics*. Oxford: Blackwell, 2001.
- TAGNIN, S. E. O. *O jeito que a gente diz: Expressões convencionais e idiomáticas*. São Paulo: Disal, 2005.
- WEINRICH, H. *Textgrammatik der deutschen Sprache*. 3ª Edição revisada. Hildesheim, Zúrique, Nova Iorque: Georg Olms Verlag, 2005.
- ZILIO, L. *Colocações especializadas e Komposita: um estudo contrastivo alemão-português na área de cardiologia*. Porto Alegre: UFRGS. Dissertação de Mestrado. PPG-LETRAS/UFRGS, 2009.
- _____. Terminologia Textual e Lingüística de Corpus: estudo em parceria. In: PERNA, C. L.; DELGADO, H. K.; FINATTO, M. J. B. (Orgs.) *Linguagens Especializadas em Corpora: modos de dizer e interfaces de pesquisa*, 2010. Disponível em: <<http://www.pucrs.br/edipucrs/linguagensespecializadasemcorpora.pdf>>.