

Gouadec e Gémar: suas contribuições para o projeto Combinatórias Léxicas Especializadas



Cleci R. Bevilacqua

cleci.bevilacqua@gmail.com

Patrícia C. R. Reuillard

patriciam@terra.com.br

Universidade Federal do Rio Grande do Sul

Gouadec e Gémar: les contributions pour le projet Combinatoires Lexicales Spécialisées

Résumé : Cet article présente le projet actuellement développé par le Projet Terminologique Cône Sud (Termisul) (2010-2015) de l'Institut de Lettres de l'Université Fédérale du Rio Grande do Sul (UFRGS). Dans son versant appliqué, le projet vise la construction d'une base de données de Combinatoires Lexicales Spécialisées (CLE) du langage juridique, normatif et scientifique à l'attention des traducteurs et rédacteurs spécialisés. Le versant théorique se penche sur le processus de cette construction pour proposer les fondements théoriques de cette pratique. Dans la présentation du référentiel théorique, l'accent est mis sur les contributions du théoricien Daniel Gouadec et du jurilinguiste Jean-Claude Gémar, ainsi que sur les étapes prévues, la méthodologie appliquée et certains des résultats obtenus jusqu'à ce jour.

Mots-clés : terminologie ; langage juridique ; langage scientifique ; combinatoires lexicales spécialisées.

Resumo: Este artigo apresenta o projeto atual desenvolvido pelo Projeto Terminológico Cone Sul (Termisul) (2010-2015), do Instituto de Letras da Universidade Federal do Rio Grande do Sul (UFRGS). Em sua vertente aplicada, o projeto tem o objetivo de construir uma base de dados de Combinatórias Léxicas Especializadas (CLEs) da linguagem legal, normativa e científica para tradutores e redatores especializados. A face teórica se debruça sobre o processo dessa construção para propor o arcabouço teórico que sustenta essa prática. Apresentaremos o referencial teórico, destacando as contribuições do teórico Daniel Gouadec e do jurilinguista Jean-Claude Gémar, as etapas previstas, a metodologia aplicada e alguns dos resultados alcançados até o momento.

Palavras-chave: terminologia; linguagem jurídica; linguagem científica; combinatórias léxicas especializadas.

Gouadec e Gémar: contributions to the Specialized Lexical Combinations project

Abstract : This paper shows the present work developed within the Projeto Terminológico Cone Sul (Termisul) (2010-2015) at the Instituto de Letras, Universidade Federal do Rio Grande do Sul (UFRGS). In its application mode, the project aims at building a database of Specialized Lexical Combinations from legal, normative and scientific language addressed to translators and specialized writers. We will present theoretical references, emphasizing contributions from theorist Daniel Gouadec and legal linguist Jean-Claude Gémar, planned stages, applied methodology, and some results already achieved.

Keywords: terminology, legal language, scientific language, specialized lexical combinations.

Introdução

Este artigo apresenta a etapa atual das pesquisas do grupo Termisul, cujo percurso começou nos anos 1990. A pesquisa teórica e aplicada que estamos desenvolvendo (2010-2015) focaliza as expressões multivocabulares específicas dos textos legais, normativos e científicos¹. O estudo concentra-se na identificação e na caracterização dessas expressões, denominadas Combinatórias Léxicas Especializadas (CLEs). Seu objetivo é refletir sobre o tema das terminologias e das fraseologias, distinguindo-se as composições lexicais especializadas, peculiares da linguagem científica, daquelas composições que têm um caráter geral e que se apresentam em textos sem associação direta com terminologias científicas. Tal reflexão busca gerar fundamentação teórica e metodológica para o oferecimento de produtos terminográficos *on-line* dirigidos principalmente a redatores técnicos e a tradutores, bem como propiciar avanços significativos para o reconhecimento automático em aplicações de Processamento da Língua Natural (PLN).

Consideramos que identificar as CLEs, explicar seu significado, conhecer seu comportamento na comunicação e sugerir alternativas para sua descrição, categorização e tratamento são temas indispensáveis para o desenvolvimento dos estudos terminológicos. Além disso, julgamos necessário o conhecimento de tais combinatórias na língua materna e nas línguas de trabalho do tradutor ou do redator para produzir textos corretos do ponto de vista linguístico e adequados do ponto de vista da convencionalidade da área especializada. Consequentemente, sua identificação e apresentação em forma de base de dados, disponibilizada *on-line*, é um dos objetivos da pesquisa. Nessa linha, o projeto inclui seis línguas: português, alemão, espanhol, francês, inglês e italiano.

Ressaltamos que, embora o público preferencial visado se constitua de tradutores e de redatores, os resultados obtidos com a pesquisa poderão ser aproveitados também por profissionais de outras áreas da Linguística e da Informática, bem como por legisladores, advogados, empresários, profissionais da comunicação, dentre outros, abrangendo vários setores da sociedade.

Por sua complexidade, os pontos de maior interesse do estudo concentram-se no próprio conceito de fraseologia ou combinatórias léxicas especializadas e em sua delimitação e, portanto, sua identificação em meio a outros tipos de combinatórias. Nesse sentido, a clara determinação do início e do fim de uma combinatória depende também da definição ou da perspectiva teórica adotada. Para os tradutores e redatores de textos especializados, no entanto, mais do que traçar limites, interessa conhecer

quais unidades lingu sticas ou coocorrentes acompanham os termos de determinada  rea e quais combinat rias s o caracter sticas de determinado dom nio discursivo. Sendo assim, nesse texto pretendemos apresentar uma vis o geral das diferentes perspectivas da fraseologia especializada, centrando-nos sobretudo nas contribui es de Daniel Gouadec e de Jean-Claude G mar, estudiosos cujas propostas de fraseologia e de equival ncia t m norteado os estudos do grupo.

2 As diferentes vis es de Fraseologia Especializada

O tema das combinat rias ou da fraseologia no  mbito da l ngua geral   bastante antigo e seu estudo tem sido objeto de pesquisas lexicol gicas e lexicogr ficas, intensificadas desde a d cada de 1980 e aprofundadas gra as   contribui o de linguistas e jurilinguistas, dicionaristas e estudiosos da l ngua natural. No entanto, sua descri o ainda   discutida, e a quest o de seu tratamento em obras de refer ncia n o foi ainda inteiramente resolvida, embora existam in meros dicion rios de combinat rias, conhecidos como dicion rios de fraseologia, de frases feitas ou de express es idiom ticas.

Por m, no  mbito das linguagens especializadas, as fraseologias dos textos especializados come aram a ser estudadas pelos linguistas e terminol gos apenas a partir dos anos 1990; em um primeiro momento, consideraram tais associa es sintagm ticas em uma perspectiva lexicol gica e lexicogr fica. Apoiados na no o de coloca es de Firth (1957) e na an lise distribucional de Harris (1991), eles trataram esse tipo de combinat rias como fraseologias da l ngua comum, seguindo as propostas de Hausmann (1990) e Mel uk (1984, 1988, 1992). Nessa vis o, tais unidades s o denominadas coloca es e s o formadas, basicamente por dois elementos: a base (elemento semanticamente aut nomo) e o colocado ou elemento coocorrente (restringido semanticamente pela base). Tais unidades s o semilexicalizadas, t m um valor semicomposicional e seu significado n o pode ser deduzido de forma independente de cada um dos elementos que a comp em, mas pelo conjunto desses elementos. Nessa  tica, s o estabelecidas certas estruturas morfossint ticas como padr o para auxiliar no seu reconhecimento, tais como: substantivo + substantivo; substantivo + adjetivo; verbo + substantivo; adjetivo + adv rbio, verbo + adv rbio. Tal concep o   adotada para o estudo das CLEs por muitos autores, entre os quais L'Homme e Bertrand (2000) e Tercedor S nchez (1998).

Para essas autoras, o objeto de estudo da fraseologia especializada s o as combinat rias formadas por termos e seus coocorrentes, entre os quais se estabelece uma restri o de sele o determinada pelas especificidades dos  mbitos especializados. Como exemplos dessas unidades, podemos citar *deletar um arquivo* (inform tica) ou *acumula o de  gua* (ecologia). Como se pode observar, n o h  uma distin o precisa entre o que se constitui em termo sintagm tico (*recursos vivos e n o-vivos*) e o que  

outro tipo de combinatória especializada (*aproveitar recursos vivos e não-vivos*). De fato, a fronteira entre o conceito de termo sintagmático e de combinatória especializada é muito tênue, e estabelecer seus limites é tarefa muitas vezes incerta e discutível.

Uma outra perspectiva em relação à concepção de fraseologia - de orientação mais terminológica - considera que as combinatórias especializadas, denominadas Unidades Fraseológicas Especializadas (UFE), são unidades sintagmáticas, formadas por dois ou mais elementos, especificando que um dos elementos é o termo e que o outro tem caráter verbal. Distingue-se das perspectivas anteriores pelas seguintes características: a) atribui um caráter verbal à unidade; b) não propõe estruturas morfossintáticas para auxiliar na sua identificação; c) não faz referência à restrição entre os elementos que compõem a unidade; d) estabelece claramente a diferença entre termos e fraseologias.

As autoras Lorente, Estopà e Bevilacqua (1998) e Bevilacqua (2004) adotam esse ponto de vista e distinguem termos sintagmáticos e UFEs. Sendo assim, *recursos vivos e não-vivos* é claramente um termo, pois possui um referente específico na área ambiental, enquanto *aproveitar* ou *aproveitamento de recursos vivos e não-vivos* é uma unidade fraseológica especializada, pois denota um processo dessa área.

Uma terceira perspectiva, distinta das anteriores e à qual dedicaremos mais atenção neste texto, é a de Gouadec (1994). O autor prefere o termo *entidades fraseológicas*, no qual inclui um conjunto flexível de combinações como palavras, grupos de palavras, termos, locuções, expressões, proposições, sem a preocupação de estabelecer categorias nem critérios morfológicos. Para ele, o objetivo da identificação e do tratamento dessas unidades é seu gerenciamento e domínio pelo redator e tradutor. Gouadec indica quatro características relevantes para seu reconhecimento: a especialização, a repetição, o risco inerente à sua manipulação e a vantagem de dominar seu uso (idem p.167-168).

Nesse contexto, Gouadec postula como critérios fundamentais para a identificação desse tipo de unidades a estereotipia, vista como um conjunto de caracteres estabilizados e recorrentes em uma área, e a frequência com que ocorrem em tal área. Concebe ainda que tais unidades se constituem de elementos invariáveis e variáveis, indicando, desse modo, a possibilidade de alterações no interior de uma mesma unidade, na qual pode ocorrer inserção e/ou supressão de um ou mais de um elemento ou alteração de ordem. A partir desses critérios (estereotipia, frequência e conformação através de invariáveis e variáveis), esses conjuntos de caracteres especializados assumem o caráter de matriz. Sua proposta é eminentemente funcional e relaciona-se diretamente aos ganhos de qualidade e produção do texto no processo tradutório. Como exemplo de matriz temos *assumir um compromisso entre [x] e [y]*, em que *assumir um compromisso entre* é a parte invariável, e *[x] e [y]*, a parte variável. As variáveis podem ser preenchidas, no caso de um contrato de aluguel, por exemplo, por *[locdor]* e *[locatário]*.

Considerando as diferentes vis es do tema e as diferentes abordagens sugeridas para trat -lo, apresentadas acima, podemos dizer que h  pelo menos tr s perspectivas te ricas relativas   fraseologias ou combinat rias especializadas, cada uma com suas especificidades: uma de perspectiva mais lexicol gica, aproximando-se da l ngua geral; outra marcadamente terminol gica, enfatizando o car ter especializado; e outra ainda de car ter pragm tico, voltada   pr tica textual e tradut ria.

3. Nossa proposta

Do quadro referencial te rico mais amplo deste trabalho, destacamos o pressuposto b sico da Teoria Comunicativa da Terminologia (Cabr , 1999, 2001), que considera a linguagem especializada como uma express o da l ngua comum usada em contexto de especialidade, compartilhando de todas as propriedades da l ngua natural ao desempenhar as fun es comunicativas pr prias da  rea tem tica. Nessa perspectiva, recorreremos   abordagem descritiva da Terminologia Textual (Bourigault; Slodzian, 1999), que analisa a linguagem especializada na sua efetiva realiza o em contexto de uso. Ao mesmo tempo, servimo-nos da Lingu stica de *Corpus*, que privilegia o estudo da l ngua em grandes conjuntos de textos aut nticos analisados por ferramentas digitais (Berber Sardinha, 2004).

Considerando tais paradigmas da Terminologia e com base sobretudo em Gouadec (1994) e Bevilacqua (1996, 2004), definimos as CLEs como unidades sintagm ticas, formadas por dois ou mais elementos, podendo conter um termo. Tais unidades s o recorrentes nas situa es de comunica o das  reas tem ticas que revelam uma prefer ncia marcante pelas especificidades e pelas conven es pr prias do idioma, da  rea e/ou do tipo de texto em que ocorrem. S o unidades semilexicalizadas n o inteiramente fixas, que n o se confundem com os termos, mas que caracterizam a especializa o de um texto. Ainda que n o partilhem do car ter de termo na estrutura conceitual de um dom nio do conhecimento, as CLEs s o de import ncia indiscut vel na comunica o do saber, porque denotam processos, atividades e a es pr prias da  rea tem tica em que ocorrem.

A partir dessas premissas, estabelecemos os seguintes crit rios para a identifica o das CLEs, baseados na proposta de Gouadec (1994):

- sintagmaticidade: devem estar formadas por mais de um elemento lingu stico;
- prototipicidade: s o espec ficas de determinadas  reas de conhecimento e caracterizam suas formas de dizer, como as combinat rias do texto legislativo *no que couber/ en tant que de besoin*;

- frequência relevante: devem ter uma frequência considerável nos âmbitos em que são utilizadas, o que auxilia a caracterizá-las como prototípicas de dada área. No entanto, esse critério pode ser flexibilizado, pois, em determinado *corpus*, uma CLE que tem uma ocorrência baixa pode ser importante na área e não deve ser excluída unicamente pela aplicação deste critério;

- consensualidade: seu uso é estabelecido pelos membros da área especializada, o que lhes confere caráter convencionalizado e de um âmbito específico. Assim, no *corpus* de legislação ambiental, encontramos *alijamento de resíduos*, mas não *exclusão de resíduos*; *revogar uma lei* e não *retirar uma lei*. Da mesma forma, quando se trata da equivalência, para a CLE de acordo com o disposto, teremos em francês *selon les dispositions* e não **selon ce qui est disposé*;

- relativa estabilidade semântica e sintática: são fixas ou semifixas, isto é, admitem pouca ou nenhuma substituição dos elementos linguísticos que as conformam, no eixo paradigmático, e inclusão de novos elementos no eixo sintagmático;

Além de definir os critérios para sua seleção, ao trabalhar com diversas línguas estrangeiras, é importante também ter presente e fundamentar-se em uma definição de equivalência. Desse modo, adotamos a definição de equivalência funcional defendida pelo jurilinguista Gémar, que estuda a tradução no âmbito do Direito (1998). Para ele, trata-se de um compromisso entre a tradução linguística (com ênfase na forma) e a tradução jurídica (com ênfase no conteúdo) (Gémar, 1998: 12). Portanto, são equivalentes as estruturas que expressam a mesma relação semântica e o mesmo efeito pragmático nos textos de partida e de chegada. Isso significa que os equivalentes em língua estrangeira devem ser buscados em textos autênticos e seus contextos, analisados, para verificar se preenchem a mesma função do texto de partida. Nessa situação, a consulta a especialistas revela-se primordial caso persistam dúvidas.

4 Corpora textuais e etapas metodológicas

As CLEs são extraídas das bases textuais já construídas pelo Grupo Termisul²:

- Base Legis: *corpus* composto por textos de legislação ambiental (Alemanha, Argentina, Brasil, Estados Unidos, França, Itália, Paraguai e Uruguai), de **Atos Internacionais** (Agenda 21, Declaração do Rio, Protocolo de Kyoto e Convenção de Estocolmo), e de **textos alinhados**³ (Declaração do Rio e Protocolo de Kyoto e Convenção de Estocolmo, sempre a partir do inglês e uma língua estrangeira).

- Base Gestamb: *corpus* composto por textos de Gestão Ambiental, tais como manuais, teses e dissertações e obras de referência em língua portuguesa⁴.

- Base Tecno-Ci ncia: *corpus* composto por textos cient ficos de Cardiologia.

Cada uma das bases constitui uma etapa do projeto, e previram-se os seguintes procedimentos metodol gicos⁵: revis o geral dos *corpora* e sistematiza o das principais caracter sticas estruturais nos diferentes idiomas; extra o de candidatos a CLEs com aux lio das ferramentas implementadas no Termisul⁶ e tamb m com aquelas dos programas AntConc (Anthony, 2008) e Multiword Expression Toolkit (Ramisch, 2012); aplica o dos crit rios para sele o das CLEs nas listas de *n-gramas* geradas pelas ferramentas; identifica o dos equivalentes nas l nguas estrangeiras; constru o e testagem da base de combinat rias em formato amig vel para disponibiliza o *on-line* (sua arquitetura ter  dois m dulos, um de trabalho e um de consulta para o usu rio); alimenta o da base com inser o das combinat rias e informa es relevantes (contextos, notas, etc.); e, por fim, publica o da base de combinat rias.

5 Resultados iniciais

Na etapa atual da pesquisa, estamos trabalhando com a coleta das CLEs da Base Legis⁷. Coletamos um conjunto significativo de combinat rias e, ap s uma an lise detalhada, foi poss vel classific -las em dois grupos: CLEs jur dicas e CLEs terminol gicas.

5.1 CLEs Jur dicas

As CLEs jur dicas s o utilizadas em textos legais de qualquer  rea do Direito e podem ter diferentes fun es:

a) indicar a es espec ficas do poder legislativo, como emitir normas, atribuir poderes e compet ncias, prescrever comportamentos, etc., como em *Revogam-se as disposi es em contr rio* ou *La pr sente loi sera ex cut e comme loi d' tat*; b) indicar remiss o interna ou intertextual, como em *de acordo com o disposto em...* ou *selon les dispositions de*; c) estabelecer restri es e/ou delimita es. Por exemplo, *Para os fins deste Decreto, entende-se por: ...*

Al m dessas fun es, as CLEs situam-se em geral em partes espec ficas da macroestrutura textual, o que permite identificar sua fun o pragm tica. Por exemplo, as CLEs *Esta lei entra em vigor na data de sua publica o e [inst ncias governamentais] sont charg s chacun en ce qui le concerne, de l'ex cution du pr sent d cret, qui sera publi  au Journal Officiel de la R publique Fran aise* ocorrem sempre nos artigos finais e t m a fun o de fechar a lei indicando a vig ncia a partir do momento de promulga o.

Outra caracter stica desse tipo de CLE   sua estrutura matricial, segundo a proposta de Gouadec, isto  , possui uma parte fixa e outra vari vel: a parte fixa   est vel e n o

sofre substituição, embora possa apresentar variações sintáticas: *para efeitos de* [texto legislativo ou parte de texto legislativo] e *para os efeitos de* [texto legislativo ou parte de texto legislativo] ou *dans le but de faire observer* [texte législatif ou partie d'un texte législatif] e *en vue de faire observer* [texte législatif ou partie d'un texte législatif]; admite variantes (sinônimos): *de acordo com a lei* tem como expressão sinonímica *nos termos da lei*; *conformément à* tem como sinônimo *selon les dispositions*; c) a parte variável é a unidade terminológica específica do âmbito do Direito: *para os efeitos do Artigo 14, para os efeitos da Lei nº 6.938/81, en vue de faire observer la présente Loi*, etc.

5.2 CLEs Terminológicas

Essas CLEs possuem as seguintes propriedades: a) caracterizam a área do Direito Ambiental, área interdisciplinar; b) referem-se às ações e/ou processos específicos do meio ambiente, mas estão circunscritas pelo Direito - a tutela; c) são formadas por verbos (*avaliar, approvisionner*) ou nomes deverbais (*avaliação, approvisionnement*) e um termo (simples ou sintagmático: *agrotóxicos; agrotóxicos, seus componentes e afins, eau; ressources naturelles*).

5.3 Criação da Base Legis

Outro resultado alcançado foi a criação da base de dados para registro das combinatorias. Essa base foi criada em linguagem MySQL e sua arquitetura se compõe de fichas terminológicas com campos inter-relacionados que permitem a utilização dos recursos de hipertexto para agilizar e dinamizar a consulta. A base possui um módulo de trabalho, utilizado pelos pesquisadores, que permite adicionar, editar ou visualizar as CLEs. Há também um módulo de consulta que estará disponível *on-line* para os usuários. Em ambos os módulos há os seguintes campos:

- CLE: campo para inserção da CLE principal;
- contexto da CLE e sua fonte: campo para inserção do contexto da CLE entrada retirado do *corpus* com indicação da fonte do documento em que foi coletado;
- outras formas: campo para inserção de CLEs variantes da CLE principal; por exemplo, a CLE *de acordo com o artigo* tem como outras formas *de acordo com o previsto no artigo* e *de acordo com o disposto no artigo*;
- contexto de outras formas: campo para inclusão do contexto das outras formas retirado do *corpus* com indicação da fonte do documento em que foi coletado;

- ver também: campo para inclusão de formas sinonímicas ou que mantenham outro tipo de relação semântica com a CLE principal; por exemplo, a CLE *de acordo com o artigo* tem como outras formas *conforme o artigo* e *de conformidade com o artigo*; cada uma dessas CLEs tem ficha própria e estão relacionadas por um *link*;

- contexto de ver também: campo para inclusão do contexto das formas sinônimas ou que tenham outra relação semântica com a CLE principal retirado do *corpus* com indicação da fonte do documento em que foi coletado;

- equivalentes nas línguas estrangeiras: campos para inclusão dos equivalentes nas línguas alemã, espanhola, francesa, inglesa e italiana; cada um dos equivalentes é um *link* que abre a ficha completa da CLE na respectiva língua;

- nota: campo para inclusão de notas de uso referentes à CLE principal.

6. Considerações finais

Como procuramos demonstrar, as contribuições de Gouadec e de Gémar foram basilares em nossa pesquisa e nortearam a busca e o estabelecimento das combinatórias léxicas especializadas da legislação ambiental. Os critérios estabelecidos por Gouadec para o estabelecimento de *entidades fraseológicas* - estereotipia, frequência e conformação através de invariáveis e variáveis - orientaram a busca de CLEs; de fato, tanto na extração semiautomática quanto na extração manual, analisamos cada combinatória selecionada de modo a responder se atendia aos critérios preestabelecidos. Essa orientação foi o que nos permitiu estabelecer o que entendemos por matriz e variáveis, como na CLE *para os efeitos de [texto legislativo ou parte de texto legislativo]*, que se compõe da matriz *para os efeitos de* e de uma parte variável, representada por um termo jurídico como portaria, lei artigo, parágrafo, etc. Além disso, essa abordagem também permitiu abrir espaço para o registro de suas variantes - *para efeito das disposições de [texto legislativo ou parte de texto legislativo]*, *para os efeitos do disposto em [texto legislativo ou parte de texto legislativo]* - e para formas que mantêm relações semânticas de sinonímia ou antonímia com a forma mais frequente, como *para os fins de [texto legislativo ou parte de texto legislativo]*.

Do mesmo modo, a noção de equivalência funcional, defendida por Gémar e que serviu de fundamento para a seleção dos equivalentes em língua estrangeira, tem reiterado nosso entendimento sobre a relevância do aspecto pragmático na busca dos equivalentes e sobre a importância da construção de um *corpus* com textos autênticos do domínio trabalhado.

Concluída esta etapa das combinatórias relativas à legislação ambiental, passaremos

àquelas da gestão ambiental, quando certamente novas questões de pesquisa demandarão reflexão e resposta.

Referências bibliográficas

- Anthony, Lawrence (2008) *AntConc* (3.2.1 w) Tokyo: Waseda University. <http://www.antlab.sci.waseda.ac.jp/software.html> [Data de consulta: 28/04/2011].
- Berber Sardinha, Toni (2004): *Linguística de Corpus*. Barueri: Manole.
- Bevilacqua, Cleci Regina (2004): *Unidades Fraseológicas Especializadas Eventivas: descripción y reglas de formación en el ámbito de la energía solar*. Tese de doutoramento. (http://www.ufrgs.br/termisul/biblioteca/teses/tese_DOUTORADO_2004_BEVILACQUA.pdf). [Data de consulta: 30/05/2010].
- Bevilacqua, C.R. et alii (2013): “Combinatórias Léxicas da Linguagem Legislativa: uma abordagem orientada pelo *corpus*”, em Murakawa, Clotilde de Almeida Azevedo; Nadin, Odair Luiz. (Eds.): *Terminologia: uma ciência interdisciplinar*. Araraquara: FCL-UNESP Laboratório Editorial/ Cultura Acadêmica Editora, p.227-244.
- Bourigault, Didier; SLODZIAN, Monique (2000): “Pour une terminologie textuelle”. *Terminologies Nouvelles*, n.19, p.29-32.
- Cabré, María Teresa (1999): *Terminologia: representación y comunicación. Una teoría de base comunicativa y otros artículos*. Barcelona: Universitat Pompeu Fabra, Institut Universitari de Lingüística Aplicada.
- Cabré, María Teresa (2001): “Sumario de principios que configuran la nueva propuesta teórica y consecuencias metodológicas”, em CABRÉ, María Teresa; FELIU, Judit (Ed.): *La terminología científico-técnica: reconocimiento, análisis y extracción de información formal y semántica*. Barcelona: Universitat Pompeu Fabra, Institut Universitari de Lingüística Aplicada, p. 17-25.
- Firth, John Rupert (1969) “Modes of meaning”, en: FIRTH, John Rupert *Papers in linguistics 1934-1951*. London: Oxford University Press, p. 190-215.
- Gémar, Jean-Claude (1998): Les enjeux de la traduction juridique. Principes et nuances. *Astti Seminar: Equivalences 1998: Traduction de textes juridiques: problèmes et méthodes* Disponível em : <http://www.tradulex.com/Bern1998/Gemar.pdf> - Acesso em 12/02/2012.
- Harris, Zelig (1991): *Theory of language and information: a mathematical approach*. Oxford: Oxford University Press.
- Gouadec, Daniel (1994): «Nature et traitement des entités phraséologiques», em *Terminologie et phraséologie. Acteurs et aménageurs: Actes de la deuxième Université d'Automne en Terminologie*. Paris: La Maison du Dictionnaire, p. 167-193.
- Hausmann, Franz Josef. (1990): “Le dictionnaire de collocations», em: Hausmann, Franz Josef et alii. *An International encyclopedia of lexicography*. Vol. 1. Berlin, New York: Walter de Gruyter, p. 1010-1019.
- L’homme, Marie-Claude; BERTRAND, Claudine (2000): “Specialized Lexical Combinations: Should they be Described as Collocations or in Terms of Selectional Restrictions?” (<http://www.ling.umontreal.ca/lhomme/docs/berlhom.pdf>) [Data de consulta 15/03/2010].
- Lorente, Mercè; BEVILACQUA, Cleci Regina.; ESTOPÀ, Rosa (2002): “El análisis de la fraseología especializada mediante elementos de la lingüística actual”, em: CORREA, Margarita (Ed.): *Atas VI Simpósio da Rede Ibero-americana de Terminologia. Terminologia, desenvolvimento e identidade nacional*. Lisboa: Ed. Colibri/ILTEC, p. 647-666.
- Mel’cuk, Igor (1984, 1988, 1992): *Dictionnaire Explicatif et combinatoire du français contemporain*. Vol. 1, 2, 3, Montréal: Les Presses de la Université de Montréal.
- Ramisch, Carlos (2012): A generic and open frame work for multiword expressions treatment: from acquisition to applications. Tese de doutoramento.

(http://www.inf.ufrgs.br/~ceramisch/download_files/thesis/ Acesso em 30/08/2012.

Tercedor Sánchez, María Isabel (1998): *La fraseología en el lenguaje biomédico: análisis desde las necesidades del traductor*. Tese de doutoramento. (<http://elies.rediris.es/elies6/>). [03/06/2010]

Notes

1 Este projeto conta com a participação das pesquisadoras: Anna Maria Becker Maciel, Cláudia Mendonça Scheeren, Cristiane Krause Kilian, Denise Sales e Sandra Dias Loguércio, além de bolsistas de iniciação científica, alunos do curso de Bacharelado em Letras-Tradução, e pesquisadores colaboradores.

2 Para mais detalhes: http://www.ufrgs.br/termisul/bases_textuais/apresentacao.php.

3 Textos alinhados: conjunto de dois ou mais textos, um original e sua(s) tradução(ões), cujos segmentos equivalentes são dispostos paralelamente.

4 Essa base possui somente textos em português e seus equivalentes serão pesquisados em outras fontes, principalmente da Internet, seguindo critérios de busca estabelecidos pela equipe.

5 Para mais detalhes sobre a metodologia de extração das CLEs, ver Bevilacqua et alii 2013.

6 Listador de palavras: lista todas as palavras do *corpus* em ordem alfabética ou de frequência; concordanciador: lista as ocorrências no texto de uma determinada palavra com seu contexto imediato; gerador de *n-gramas*: sequências de palavras que se repetem no *corpus*, que podem ter diversas extensões; gerador de *clusters*: agrupamento de palavras ao redor de uma palavra-chave. Para mais detalhes sobre essas ferramentas, ver

http://www6.ufrgs.br/termisul/limbo/ferramentas_novo/Seleciona.php.

7 Projeto contemplado com recursos no edital 18/2012 do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).