

ANÁLISE DE ÍNDICES DE COMPLEXIDADE TEXTUAL EM DOIS TEXTOS DE PEDIATRIA E SUAS TRADUÇÕES PARA O INGLÊS: UM ESTUDO EXPLORATÓRIO

Bianca Franco Pasqualini (Mestranda em Linguística, UFRGS)

bianca.pasqualini@gmail.com

RESUMO: Este trabalho propõe-se a analisar índices de complexidade textual de uma amostra de dois textos técnico-científicos da área de Pediatria e suas respectivas traduções para o inglês. Para isso, é traçado um breve panorama dos Estudos Terminológicos, seguido da exposição das ferramentas utilizadas: o Coh-Metrix e o Coh-Metrix-Port. Após seleção das métricas a serem analisadas e de uma exposição sobre algumas delas, os resultados são apresentados e discutidos. Em seguida, elencam-se algumas perspectivas, que apontam na direção da construção de uma ferramenta de análise de complexidade textual para tradutores baseada em princípios da Análise Multidimensional.

Palavras-chave: terminologia, complexidade textual, tradução, análise multidimensional

INTRODUÇÃO

O aprofundamento dos Estudos de Tradução e da Terminologia como área do conhecimento vem oferecendo cada vez mais ao tradutor bases teóricas e ferramentas para aprimorar traduções de textos técnico-científicos. Hoje, o tradutor conta não só com obras terminográficas e bancos de dados de linguagens de especialidade, como

também tem à disposição novas perspectivas sobre o texto que produz – perspectivas que ampliam visões focadas exclusivamente no termo, passando também a considerar o ambiente textual em que tais termos ocorrem.

Neste breve estudo exploratório, usamos uma amostra composta por dois textos da área de Pediatria e suas respectivas traduções para o inglês e, como textos contrastivos, dois contos da literatura brasileira e suas traduções também para o inglês. Alguns índices de complexidade desses textos foram analisados através das ferramentas Coh-Metrix e Coh-Metrix-Port. Tais índices englobam aspectos lexicais, sintáticos e semânticos, além de índices de legibilidade como, por exemplo, o índice Flesch. Antes de discutir as ferramentas e os resultados, são traçados comentários sucintos sobre o desenvolvimento da Terminologia a fim de contextualizar o escopo do trabalho.

A TERMINOLOGIA E SUAS VERTENTES

Teoria Geral da Terminologia

A Terminologia moderna surge a partir dos anos 30 e funda-se no fim da década de 70 com a publicação póstuma da obra *Introducción a la Teoria General de la Terminologia y a la Lexicografia Terminológica* (WÜSTER, 1998), cuja versão original foi publicada na Alemanha em 1979. A partir dessa publicação, a Terminologia alcança estatuto de disciplina e ciência, com objeto próprio de estudo.

Nessa visão fundadora, os termos são designações de conhecimento científico, e não elementos naturais da língua. Assim, a fim de padronizar o uso da terminologia, a motivação central era a de controlar a linguagem, rotulando e etiquetando conceitos no intuito de evitar as ambiguidades comunicativas do léxico geral entre especialistas. Tais

conceitos divergem da noção saussuriana de significado, uma vez que consideram os conceitos científicos como atemporais, paradigmáticos e universais, revelando, além disso, uma teoria cujos princípios epistemológicos fundamentam-se na dissociação entre pensamento e linguagem (KRIEGER e FINATTO, 2004, p. 33).

Vê-se, portanto, que Teoria Geral da Terminologia (TGT) preocupa-se com a sistematização dos conceitos técnico-científicos e com a padronização terminológica da transmissão desses conceitos (tanto numa mesma língua como em tradução), num esforço para eliminar as variações e polissemias intrínsecas aos signos linguísticos.

A TGT trata os conceitos como anteriores à sua expressão na linguagem e, por consequência, o léxico dos sistemas conceituais como nomenclatura ou taxonomia. Suas bases são os parâmetros de ciência que incluem a busca da objetividade, da precisão e da não ambiguidade com raízes na filosofia analítica e no neopositivismo. Postula um ideal de linguagem unívoca, homogênea e universal (que se distingue da linguagem comum) e concebe os conceitos como constructos mentais anteriores e independentes da linguagem. Ao privilegiar esses aspectos, a TGT não leva em conta uma das funções primordiais da linguagem que é a de ser instrumento de comunicação.

No entanto, releituras recentes da obra wüsteriana indicam que ele também reconheceu a interface da Terminologia com a Linguística ao concebê-la como um ramo da Linguística Aplicada, conforme aponta Krieger (2001, p. 49): “Tal reconhecimento reflete-se no quadro de apreensão dos mecanismos de funcionamento dos termos, em especial, quando faz referência às noções de monovalência ou de univocidade”.

Socioterminologia

Contrapondo-se aos postulados da TGT, a Socioterminologia, corrente teórica divulgada a partir dos anos 90, sustenta a importância da variação terminológica como representativa da identidade linguística e cultural das terminologias. Os principais autores vinculados a essa escola são Boulanger, Gambier e Gaudin, com desenvolvimento sobretudo em países bilíngues como o Canadá.

A Socioterminologia, ao contrário da TGT, discute o papel dos termos em uso em seu contexto social e propõe que os termos devem ser analisados *in vivo*, e não *in vitro* (TEIXEIRA, 2008). Nessa perspectiva terminológica, os termos são antes de mais nada signos linguísticos, invertendo a lógica da visão tradicional da Terminologia de que os conceitos se sobrepõem à forma e ampliando as restrições normativas determinadas pela ideia de que é possível eliminar ambiguidades discursivas nos discursos técnico-especializados. Assim, obras terminográficas de cunho socioterminológico teriam a missão de abarcar as diferentes formas associadas a um conceito de acordo com as comunidades que as utilizam e com a circulação dessas formas na prática discursiva dos usuários.

Teoria Comunicativa da Terminologia

A Teoria Comunicativa da Terminologia (TCT) também funda-se a partir de críticas à TGT ao afirmar, tal como em boa medida fez a perspectiva socioterminológica, que os termos estão inseridos numa língua natural. A força terminológica de um item lexical, para a TCT, não se dá sem que antes seu valor linguístico seja avaliado num contexto de comunicação especializada, ou seja, “o

conteúdo de um termo não é fixo, mas relativo, variando conforme o cenário comunicativo em que se inscreve” (KRIEGER e FINATTO, 2004, p. 35). Nesse sentido, Cabré (*apud* TEIXEIRA, 2008 p. 60) destaca as seguintes características do léxico especializado:

- Sua poliedricidade denominativa, cognitiva e funcional.
- Sua funções representativa e comunicativa.
- Seus elementos operantes: a linguagem é real, e não ideal, e ocorre *in vivo*.
- Suas aplicações diversas conforme a situação comunicativa.

Teixeira (2008) também ressalta o fato de que a TCT, a partir dessa concepção multifacetada proposta por Cabré, “é uma teoria linguística de base cognitiva e propósito comunicativo” (p. 61). Além disso, a TCT tem no uso de *corpora* uma fonte de identificação de “unidades de conhecimento especializado”, numa interface com a Linguística de *Corpus*, com o propósito de descrever as características gerais dos discursos especializados e produzir materiais terminográficos a partir das análises levantadas em tais *corpora*.

A visão de “termo” contemplada na TCT parte da ideia de que a informação conforma-se em Unidades de Significação Especializadas, cujos eventuais nós na estrutura de conhecimento de uma área do conhecimento constituem os termos (TEIXEIRA 2008, p. 62), que são unidades de caráter denominativo e de valor referencial. Outro aspecto importante da TCT diz respeito ao fato de que, nessa concepção, os termos passam a ser usados numa área, o que indica que seu sentido se estabelece no texto e que o trabalho do terminólogo se dá numa via dirigida do termo

para o conceito (semasiológico), e não do conceito para o termo (onomasiológico, tal como postula a TGT).

Em outras palavras e fazendo uma síntese do exposto sobre a TCT, a função dos termos aqui é dupla: eles representam e transferem conhecimentos especializados em graus e modos distintos e em situações diversas, e as unidades terminológicas de carácter poliédrico desempenham funções cognitiva, social, comunicativa e linguística.

Teoria Sociocognitiva da Terminologia

Para a Teoria Sociocognitiva da Terminologia (TST), inaugurada por Rita Temmerman (KRIEGER e FINATTO, 2004; TEIXEIRA, 2008, p. 58), o enfoque terminológico deve ser baseado em modelos cognitivos e culturais, num rompimento epistemológico com a visão racionalista da TGT. Aqui, o objeto de análise é o termo como unidade interpretativa do discurso, os quais se encontram organizados em estruturas categoriais prototípicas. Essas categorias possuem um núcleo e uma estrutura, mas estão em transformação contínua conforme o conhecimento se reorganiza e se amplia.

A TST é baseada na semântica cognitiva e questiona a normatização em detrimento de uma descrição realista dos termos tal como ocorrem nos textos. Para Temmerman (*apud* TEIXEIRA, 2008, p. 59), a definição de conceito da TGT é restritiva, já que os conceitos existem de forma subjetiva. A compreensão do mundo se dá a partir de modelos cognitivos nos quais diferentes unidades de compreensão ou entendimento, estruturadas de forma prototípica, relacionam-se. A base para a delimitação do conteúdo é, nessa visão, o texto no qual está inserido o termo, do que

decorre que o conceito não é universal nem imutável, "mas a expressão de um conjunto de elementos de natureza linguística que se consubstanciam num texto que possui não apenas uma dimensão linguística, mas também pragmática, discursiva e comunicativa" (BARROS, 2006, p. 23).

A abordagem de Temmerman destaca as relações entre a Terminologia e a linguística textual, enfatizando a ideia de que os termos não podem ser compreendidos fora dos textos, que são "conjuntos dinâmicos de elementos linguísticos, pragmáticos, discursivos e comunicativos" (BARROS, 2006, p. 23-25).

Terminologia Textual

Como bem sintetiza Azeredo (2007, p. 31), as linguagens especializadas representam "um todo que especifica um *modus dicendi* típico de uma determinada área de especialidade", o que envolve mais do que simplesmente a presença de termos. Assim, para a Terminologia Textual (TT), "o texto deve ser o principal foco de estudo, pois é através dele como um todo que se poderá explicar a preferência por determinados recursos linguísticos em relação a outros" (id., p. 32). Em outras palavras, o foco da TT extrapola os limites do termo e abarca também os padrões e recorrências de estruturas linguísticas dos textos produzidos nas diferentes áreas de especialidade. Logo, a Terminologia Textual e a Linguística de *Corpus* travam um diálogo bastante próximo, pois uma fornece material e base teórica à outra no que se refere ao estudo textual do discurso especializado.

Além disso, uma perspectiva que contempla a textualidade do discurso especializado dialoga também com os Estudos de Tradução, na medida em que oferece

ao tradutor de textos técnico-científicos ferramentas que o auxiliam na tradução do texto como um todo, e não somente no que se refere aos termos encontrados em tais textos.

Conforme afirma Finatto: “Em diferentes modalidades ou tipologias, o texto especializado é construído por um enunciador que se particulariza ao mesmo tempo em que segue e prossegue um *continuum* de conhecimentos e de práticas discursivas, abrigando inclusive toda uma série de intertextualidades. [...] Primeiro há um todo do texto; depois, um modo de dizer que o faz específico. (2007, p. 449)

Conhecer essas especificidades é fundamental ao tradutor e ao revisor de textos, tendo em vista que a precisão terminológica não é garantia de uma boa tradução ou de uma boa revisão, respectivamente.

De textos com termos à complexidade textual: Coh-Metrix e Coh-Metrix-Port

Com a ampliação do terreno de estudo da Terminologia, que começa a adentrar no espaço do texto, tecnologias de áreas como a Linguística Computacional, a Psicolinguística e o Processamento de Língua Natural passam a integrar as ferramentas disponíveis a terminólogos, tradutores e pesquisadores. Uma delas é o Coh-Metrix, que significa “*cohesion metrics*”, elaborada por pesquisadores da Universidade de Memphis, nos Estados Unidos. O propósito dessa ferramenta é calcular índices de coesão e coerência textual num amplo espectro de medidas lexicais, sintáticas, semânticas e referenciais a fim de indicar a adequação de um texto a seu público-alvo (a “demanda cognitiva” e a legibilidade do texto) e de apontar problemas textuais de ordem estrutural. Até o momento, mais de quinhentas métricas estão disponíveis na versão

restrita do Coh-Metrix. Dessas quinhentas, sessenta estão disponíveis na versão gratuita *on-line* no *site* do projeto.

A partir do Coh-Metrix em inglês, uma iniciativa de adaptação para o português brasileiro das sessenta métricas gratuitas surgiu no Projeto PorSimples, cujo objetivo era o de identificar índices de complexidade textual para simplificação de textos e facilitação do acesso à informação a analfabetos funcionais e pessoas com deficiências cognitivas. O nome da ferramenta em português é Coh-Metrix-Port e está disponível no *site* do PorSimples. É importante ressaltar que até o momento apenas 35 das sessenta métricas foram adaptadas.

Essas ferramentas, ainda que não tenham sido criadas com o intuito de serem usadas na análise e na comparação de textos especializados ou de traduções, abrem um universo de possibilidades para os Estudos de Tradução, para a Terminologia e para a Linguística Aplicada, pois acrescentam uma dimensão ainda pouco explorada por essas disciplinas: a dimensão da complexidade do texto.

MATERIAIS E MÉTODOS

Com o intuito de fazer um estudo exploratório de índices de complexidade de dois textos da área de Pediatria¹ e suas respectivas traduções, as ferramentas Coh-Metrix (doravante CM) e Coh-Metrix-Port (doravante CMP) foram usadas de forma contrastiva. Além disso, dois textos literários brasileiros e suas traduções para o inglês britânico² também foram analisados para comparação com os resultados dos textos de Pediatria.

Num primeiro momento, selecionamos as métricas a serem analisadas, reduzindo-as de 35 para treze e englobando todas as categorias de análise (lexicais, sintáticas e semânticas, tendo em vista que a categoria referencial ainda está em construção). Então, ao compararmos os índices do CM com os do CMP, verificamos que nem todos têm grandezas equiparáveis, como mostra a figura 1 a seguir:

ÍNDICES	COH-METRIX-PORT	EQUIVALÊNCIA DOS ÍNDICES	COH-METRIX-ING
Índices Lexicais	Índice Flesch	EQUIVALENTE	Flesch Reading Ease
	Número de palavras	EQUIVALENTE	Number of words
	Incidência de palavras de conteúdo	NÃO EQUIVALENTE	Concreteness content words
	Frequências	TALVEZ	Raw frequency
	Mínimo frequências	TALVEZ	Min. Raw frequency
Índices Sintáticos	Operadores lógicos	EQUIVALENTE	Logic Operators
	Palavras antes de verbos principais	EQUIVALENTE	Words before main verb
	Types/token	EQUIVALENTE	Types/token
	Incidência de Sintagmas Nominais	EQUIVALENTE	NP incidence
	Incidência de conectivos	EQUIVALENTE	All connectives
Índices Referenciais e Semânticos	Anáforas	EQUIVALENTE	Anaphor reference
	Sobreposição de argumentos adjacentes	EQUIVALENTE	Adjacent argument overlap
	Sobreposição de argumentos	EQUIVALENTE	Argument overlap

Figura 1 – Lista de equivalências entre os índices do Coh-Metrix e do Coh-Metrix-Port.

No caso do índice “Incidência de palavras de conteúdo”, a métrica é, para fins contrastivos, não comparável ao índice “Concreteness content words”, pois este usa um banco de dados de concretude de palavras³ de conteúdo ainda indisponível em português. Os índices de frequências também não são equiparáveis, tendo em vista que o CM usa o banco de dados CELEX e o CMP usa o Banco de Português. Assim, as métricas passíveis de análise foram as seguintes:

Índices lexicais	Índice Flesch; Número de palavras
Índices sintáticos	Palavras antes de verbos principais; <i>Types/token ratio</i> ; Sintagmas nominais; Conectivos (causais, aditivos, temporais, lógicos) Operadores lógicos
Índices semânticos	Anáforas; Sobreposição de argumentos adjacentes; Sobreposição de argumentos

Figura 2. Índices selecionados.

Dentre os índices semânticos, escolhemos apenas as anáforas em função do tamanho reduzido deste estudo. A lista de operadores lógicos e de conectivos está no Anexo 1.

Antes de prosseguir, cabe um esclarecimento acerca do **Índice Flesch** (CMP) e **Flesch Reading Easy** (CM). Este índice foi proposto na década de 40 pelo austríaco Rudolph Flesch, que havia fugido da Europa nazista durante a guerra e se repatriado nos Estados Unidos. Sendo estrangeiro, acreditava que o *plain English*, ou seja, um inglês claro, objetivo, simples e sem ambiguidades, deveria ser o inglês usado em documentos oficiais para fácil compreensão de todos os cidadãos. A fórmula inicial de Flesch foi atualizada no CM e adaptada em 1996 para o português brasileiro por pesquisadores do Instituto de Ciências Matemáticas de São Carlos⁴. A fórmula adaptada é a seguinte:

$$ILF = 248.835 - (1.015 \times ASL) - (84.6 \times ASW),$$

onde ASL é o número de palavras dividido pelo número de sentenças e ASW é o número de sílabas dividido pelo número de palavras. Os resultados são agrupados em quatro categorias de complexidade, conforme mostra a Figura 3.

Índices entre 75 - 100	Muito fáceis Textos adequados para leitores com nível de escolaridade até a quarta série do ensino fundamental
Índices entre 50 - 75	Fáceis Textos adequados a alunos com escolaridade até a oitava série do ensino fundamental
Índices entre 25 - 50	Difíceis Textos adequados para alunos cursando o ensino médio ou universitário
Índices entre 0 – 25	Muitos difíceis Textos adequados apenas para áreas acadêmicas específicas

Figura 3. Índice Flesch.

Cabe também esclarecer que o limite operacional do CM e do CMP é de 15 mil caracteres, o que nos obrigou a limitar a análise de cada um dos textos somente até este número ou apenas até seções correspondentes (como, por exemplo, as introduções).

RESULTADOS

Os resultados foram os seguintes:

Índices	Jped – Texto 1		Jped – Texto 2	
	Português	Inglês	Português	Inglês
Número de caracteres	7.084	6.688	7.463	7.842
<i>Types/token ratio</i>	0,63	0,64	0,62	0,619
Número de palavras	1068	1044	1171	1277
Operadores lógicos	45,88	63,218	37,57	43,853
Incidência de sintagmas nominais	304,307	283,525	298,035	245,106

Conectivos	71,16	86,207	59,77	67,345
Anáforas	0,0	0,046	0,0	0,063
Índice Flesh	15,48	30,391	17,48	27,722

Figura 4 – Resultados dos textos do JPed.

Índices	A promessa/ Light of my life		Cantiga de esponsais/ Wedding song	
	Português	Inglês	Português	Inglês
Número de caracteres	6.587	6.500	8.066	8.844
Types/token ratio	0,794788	0,762	0,61385	0,576
Número de palavras	1108	1183	1435	1653
Operadores lógicos	37,9061	49,028	55,0523	37,508
Incidência de sintagmas nominais	284,2960	285,714	230,62	300,06
Conectivos	74,0072	79,459	86,4111	61,101
Anáforas	0,1333	0,224	0,391304	0,219
Índice Flesch	49,0490	65,06	66,87	78,623

Figura 5 – Resultados dos textos literários.

Quanto ao **índice *types/token ratio***, vê-se que a variação entre os textos do JPed foi bastante baixa ou quase nula. Já nos textos literários ela é mais expressiva, e mostra uma variação menor nas traduções para o inglês, ou seja, o vocabulário foi mais repetido.

O **número de palavras** dos dois textos de Pediatria analisados não mostra um padrão, pois um deles é traduzido com um maior número de palavras e o outro com menos. Já as traduções dos textos literários apresentam ambas maior número de palavras.

Os textos de Pediatria traduzidos apresentam um número expressivamente maior de **operadores lógicos** (ver Anexo 1), enquanto uma das traduções dos textos literários mostra um número maior de operadores lógicos e a outra não.

As traduções dos textos do JPed mostram ambas uma incidência menor de **sintagmas nominais**. Já as traduções literárias apresentam incidência equivalente no caso do texto 1 e incidência bem maior de sintagmas nominais no caso do texto 2.

No que se refere aos **conectivos**, as traduções do JPed têm uma incidência maior dos que os textos-fonte em ambas as traduções. As traduções literárias têm incidência maior no caso do texto 1 e menor no caso do texto 2.

A **incidência de anáforas** mostrou-se superior em todas as traduções, com exceção do texto Wedding Song.

O **índice Flesch** foi o índice que apresentou maior variação: todas as traduções, tanto as dos textos de Pediatria quanto os textos literários, têm índices Flesch mais altos, ou seja, sugerem uma legibilidade maior.

Uma vez que a amostra é bastante reduzida, é inviável fazer generalizações acerca dos textos analisados no que diz respeito ao gênero, ao par de línguas e à direção da tradução (e se feita por nativo ou não). No entanto, podemos perceber que, mesmo com uma amostragem reduzida e um número reduzido também de índices, é possível, ao menos na comparação dos textos-fonte com seus textos traduzidos, fazer um levantamento estatístico de elementos coesivos que de outra forma não seriam nem visíveis nem calculáveis. Um exemplo disso seria a proporção *types/token*, a incidência de sintagmas nominais e a incidência de anáforas consideradas em relação umas com as outras num texto especializado, pois uma variação pequena de vocabulário (um índice

types/token baixo), uma grande incidência de sintagmas e uma baixa presença de anáforas indica um elevado grau de repetição vocabular, o que, para um texto técnico-científico, talvez seja um indício de *legibilidade* e de densidade terminológica. Esses dados podem ser bastante úteis ao tradutor, que pode traçar estratégias tradutórias bem mais amplas e que contemplem mais do que o aspecto terminológico do texto.

PERSPECTIVAS E CONCLUSÕES

Para que uma ferramenta de análise de métricas de complexidade textual voltada para tradutores venha a ser criada, uma diversidade de variáveis precisa ser levada em conta, como, por exemplo, o par de línguas envolvido, a proficiência do tradutor, o público-alvo, os gêneros textuais e aspectos diacrônicos das línguas envolvidas. Além disso, as amostras precisam ser representativas e criteriosas para validar os resultados.

Há também a necessidade de avaliar a perspectiva da Análise Multidimensional proposta por Douglas Biber⁵ a fim de conciliá-la com uma proposta de análise de complexidade que seja também composta em múltiplos níveis com múltiplas relações entre si, as quais modificam umas às outras. Em outras palavras, após este estudo exploratório ficamos com a certeza de que um índice, por si só, não indica o grau de complexidade de um texto. É a relação de um índice com outros que a complexidade pode ser calculada.

Este breve estudo exploratório, ainda que pouco conclusivo em termos de resultados objetivos, revela um potencial de análise textual bastante promissor para tradutores, terminólogos e pesquisadores. Os frutos de uma pesquisa nesta direção serão

úteis na prática dos profissionais do texto, que se tornam, cada vez mais, “engenheiros do texto”.

ABSTRACT: This study analyzes textual complexity indexes of two scientific text samples (collected from a scientific journal on Pediatrics) and their translations from Brazilian Portuguese to standard English. After a brief overview of main theoretical concepts in Terminology Studies, the tools used in this study are explained and the results are presented and discussed. Lastly, some perspectives are pointed out, suggesting a text complexity tool designed specifically for translators, based on principles of Multidimensional Analysis.

Keywords: Terminology; Text complexity; Translation; Multidimensional Analysis

REFERÊNCIAS

- AZEREDO, Suzana de. **Expressões anunciadoras de paráfrase em manuais acadêmicos de Química: um estudo baseado em corpus**. Dissertação. Instituto de Letras, UFRGS, Porto Alegre, 2007.
- BAKER, Mona. “Corpus Linguistics and Translation Studies – Implications and Applications.” In: **Text and Technology: In Honour of John Sinclair**, por M. BAKER, M. G. FRANCIS e E. TOGNINI-BORELLI. Amsterdam & Philadelphia: John Benjamins, 1993.
- BARROS, Lídia Almeida. “Aspectos epistemológicos e perspectivas científicas da terminologia.” **Cienc. Cult. [online]** v. 58, n. 2 (2006).
- COH-METRIX 2.0. Universidade de Memphis. 2006. <http://cohmetrix.memphis.edu/cohmetrixpr/index.html> (acesso em agosto de 2010).
- FINATTO, Maria José B. “Exploração terminológica com apoio informatizado: perspectivas, desafios e limites.” In: **As ciências do léxico: lexicologia, lexicografia,**

terminologia. ISQUERDO, Aparecida Negri; ALVES, Ieda Maria. São Paulo: Humanitas, 2007, pp. 447-458.

ISQUERDO, Aparecida Negri; ALVES, Ieda Maria. **As ciências do léxico: lexicologia, lexicografia, terminologia.** Vol. III. São Paulo: Humanitas, 2007.

KRIEGER, Maria da Graça; Maria José B. FINATTO. **Introdução à Terminologia: teoria e prática.** São Paulo: Contexto, 2004.

KRIEGER, Maria da Graça. Do reconhecimento de terminologias: entre o linguístico e o textual. In: **As ciências do léxico.** ISQUIERDO, Aparecida Negri; KRIEGER, Maria da Graça. Campo Grande: Ed. UFMS, 2004.

KRIEGER, Maria da Graça, e Anna Maria Becker MACIEL. **Temas de Terminologia.** Porto Alegre/São Paulo: Ed. da Universidade/UFRGS/ Humanitas/USP, 2001.

REYS, Alain. A terminologia entre a experiência da realidade e o comando dos signos. In: **As ciências do léxico: lexicologia, lexicografia, terminologia.** Vol. III. ISQUERDO, Aparecida Negri; ALVES, Ieda Maria. São Paulo: Humanitas, 2007.

SCARTON, Carolina, e Daniel Machado ALMEIDA. **Coh-Metrix-Port.** 2009. <http://caravelas.icmc.usp.br:3000/> (acesso em agosto de 2010).

SCARTON, Carolina, e Sandra Maria ALUÍSIO. “Análise da Inteligibilidade de textos via ferramentas de Processamento de Língua Natural: adaptando as métricas do Coh-Metrix para o português.” **LinguaMática** 2 (2010).

TEIXEIRA, Elisa Duarte. **A Linguística de Corpus a serviço do tradutor: proposta de um dicionário de culinária voltado para a produção textual.** Tese, Departamento de Letras Modernas, USP, São Paulo: USP, 2008.

WÜSTER, E. **Introducción a la teoria general de la terminologia y a la lexicografia terminológica.** Trad. Anne-Céceli Nokerman. Barcelona: Universitat Pompeu Fabra, 1998.

Anexo 1

Operadores lógicos		
e	nada	a menos que

ou	nunca	sem que
não	jamais	a não ser que
nem	caso	salvo se
se	desde que	exceto se
nenhum	contanto que	então é porque
nenhuma	uma vez que	fosse... fosse

Figura 6 – Operadores lógicos analisados pelo CMP.

Logical connectors			
a consequence of	consequently	if	once again
actually	correspondingly	in (short/brief)	Only if
all in all	due to	in actual fact	provided that
also	enable	in any (case/event)	purpose (of/for) which
anyway	essentially then	in case	pursuant to
arise from	even then	in conclusion	secondly ;
arise out of	finally	in fact	similarly
as a consequence	first ; (next/second)	in order that	since
as a final point	first ; then	in other words	so
as a result	follow that	in sum	summarizing
as if	For	incidentally	summing up
as well	for (the/these/that) purpose	instead	That is (to say)*
at least	for example	it follow that	the consequence of
at this point	for instance	likewise	Then
because	fortunately	moreover	Then again

besides	further	Next	therefore
cause	furthermore	on (the)* condition that	thereupon
conditional upon	hence	on (the)* one hand	Thus

Figura 7 – Alguns dos operadores lógicos analisados pelo Coh-Metrix.

¹ Gentilmente cedidos por Cybelle Alle; os textos foram retirados de um *corpus* compilado da revista JPed do ano de 1995.

² Gentilmente cedidas por Francis Johnson. Os textos são: “Cantiga de esponsais”, de Machado de Assis, e sua tradução; e “A promessa”, de Humberto de Campos, e sua tradução.

³ A base de dados é a MCR Psycholinguistic Database; ver http://www.psych.rl.ac.uk/MRC_Psych_Db.html

⁴ São eles: Tereza B. F. Martins, Claudete M. Ghiraldelo, M. Graças V. Nunes e O. N. Oliveira Jr. Para consultar a pesquisa, ver *Readability formulas applied to textbooks in brazilian portuguese*, disponível no site do Projeto PorSimples: <http://caravelas.icmc.usp.br/wiki/index.php/Principal>.

⁵ Ver: BIBER, Douglas. *Variation across speech and writing*. CUP, 1988.