

UNIVERSIDADE FEDERAL DE SANTA CATARINA  
CENTRO DE COMUNICAÇÃO E EXPRESSÃO  
PROGRAMA DE PÓS-GRADUAÇÃO EM ESTUDOS DA TRADUÇÃO

**ESTELA MARIA FAUSTINO DE CARVALHO**

**METODOLOGIA DE CONSTRUÇÃO DE UM  
GLOSSÁRIO BILÍNGÜE COM BASE EM UM *CORPUS* DE  
DOMÍNIO TÉCNICO**

Florianópolis

(2007)

**ESTELA MARIA FAUSTINO DE CARVALHO**

**METODOLOGIA DE CONSTRUÇÃO DE UM  
GLOSSÁRIO BILÍNGÜE COM BASE EM UM *CORPUS* DE  
DOMÍNIO TÉCNICO**

Dissertação de mestrado apresentada ao Programa de Pós-Graduação em Estudos da Tradução, da Universidade Federal de Santa Catarina, como parte dos requisitos para a obtenção do título de Mestre em Estudos da Tradução.

Área de concentração: Lexicografia, Tradução e Ensino de Línguas.

**ORIENTADOR: Prof. Dr. Marco Antonio E. da Rocha.**

Florianópolis

2007

ESTELA MARIA FAUSTINO DE CARVALHO

**METODOLOGIA DE CONSTRUÇÃO DE UM  
GLOSSÁRIO BILÍNGÜE COM BASE EM UM *CORPUS* DE  
DOMÍNIO TÉCNICO**

Esta dissertação foi julgada adequada para a obtenção do título de Mestre em Estudos da Tradução e aprovada em sua forma final pelo Orientador e pela Banca Examinadora.

Orientador: \_\_\_\_\_

Prof. Dr. Marco Antonio Esteves da Rocha, UFSC.  
Doutor pela Universidade de Sussex, Inglaterra.

Banca Examinadora:

Profª. Dra. Anna Maria Becker Maciel, UFRGS.  
Doutora pela Universidade Federal do Rio Grande do Sul, Brasil.

Prof. Dr. Philippe René Marie Humblé, UFSC  
Doutor pela Universidade Católica de Lovaina, Bélgica.

Prof. Dr. Markus Weininger, UFSC  
Doutor pela Universidade Federal de Santa Catarina, Brasil

Coordenador da P.G.E.T.: \_\_\_\_\_  
Profª Dra. Marie-Hélène Catherine Torres

Florianópolis, maio de 2007.

## DEDICATÓRIA

Dedico este trabalho em especial ao meu marido, Orlando, pela inspiração em relação ao tema, pelo suporte e consultoria técnica.

Gostaria também de dedicar este trabalho, de forma simbólica, a todos os estudantes universitários do Brasil. Independentemente do momento político e econômico pelo qual estejamos passando, é importante acreditar em si mesmo, no seu próprio potencial. Programas de governo e projetos educacionais têm pouco ou nenhum efeito se cada um não fizer a sua parte.

Por último, mas não menos importante, gostaria de dedicar o resultado de todo o meu esforço aos meus pais que, apesar de sua pouca instrução formal, usaram sua sabedoria e experiência de vida para me ensinar que o estudo é um *bem* extremamente precioso, que felizmente não pode ser roubado, queimado ou desapropriado.

## AGRADECIMENTOS

Ao Programa de Pós-Graduação em Estudos da Tradução, P.G.E.T., em especial ao meu orientador, Prof. Dr. Marco Rocha, pela oportunidade de realização de trabalhos em minha área de pesquisa.

Ao corpo docente do P.G.E.T., pela dedicação e excelente contribuição através dos cursos ministrados na área.

Aos colegas da P.G.E.T. pela oportunidade de compartilhar conhecimentos e experiências, especialmente à Norma Andrade da Silva, pelo seu auxílio nas tarefas desenvolvidas durante o curso e apoio na revisão da minha apresentação.

Às empresas Ford Motor Company Brasil Ltda., Nissan do Brasil Automóveis Ltda, Renault do Brasil S.A. e Troller Veículos Especiais S.A., pela cessão dos direitos de uso privado dos textos constantes nos manuais destinados aos usuários dos seus veículos.

Não é possível nomear todos os profissionais que indiretamente contribuíram para o meu aprimoramento durante mais de vinte anos de carreira profissional, todavia mentalmente agradeço a todos eles.

"Jamais considere seus estudos como uma obrigação, mas como uma oportunidade invejável para aprender a conhecer a influência libertadora da beleza do reino do espírito, para seu próprio prazer pessoal e para proveito da comunidade à qual seu futuro trabalho pertencer".

Albert Einstein  
(1879 – 1955)

## RESUMO

Esta dissertação foi produzida com o objetivo de apresentar uma metodologia para a elaboração de um glossário bilíngüe com base em um *corpus* de domínio técnico. O objetivo do glossário, produto desta metodologia, é dar suporte à tradução especializada, ao ensino de línguas, ao ensino de disciplinas técnicas, à indústria e o contexto empresarial em geral, assim como ao jornalismo técnico. São apresentados os fundamentos teóricos, a metodologia propriamente dita e exemplos de extração de dados de um *corpus* do domínio técnico-automobilístico.

**Palavras-chave:** Tradução. Terminologia. Terminografia. Linguística de *Corpus*.  
**Lexicografia. Indústria Automobilística.**

## ABSTRACT

This document aims to describe a methodology for the construction of a bilingual glossary based on a domain-specific *corpus*. The bilingual glossary, result of said methodology, would have the purpose of giving support to specialized translation, language teaching, to the industry and business context in general, as well as to technical journalism. The theoretical basics, the methodology itself and examples of extraction of data from an automobile domain *corpus* are presented.

**Keywords:** Translation. Terminology. Terminography. *Corpus* Linguistics. Lexicography. Automotive Industry.



## RESUME

Ce document a pour objet de décrire une méthodologie pour la construction d'un glossaire bilingue, basé sur un *corpus* de domaine spécifique. Le glossaire bilingue, résultat de cette méthodologie, a l'objectif de servir d'appui à la traduction spécialisée, à l'enseignement de langues, à l'industrie et au contexte d'entreprise concernée, ainsi qu'au journalisme technique. Les fondamentaux théoriques, la méthodologie proprement dite et les exemples d'extraction de données d'un *corpus* du domaine technique automobile y sont présentés.

**Mots-clés : Traduction. Terminologie. Terminographie. Linguistique de *Corpus*.  
Lexicographie. Industrie Automobile.**

## SUMÁRIO

	Página
<b>1. INTRODUÇÃO</b>	14
1.1. Lingüística de <i>Corpus</i>	14
1.2. Terminologia baseada em <i>corpus</i>	15
1.3. Pesquisa em um <i>corpus</i> de conteúdo técnico	16
<b>2. REVISÃO DA LITERATURA</b>	24
<b>3. JUSTIFICATIVA</b>	24
3.1. Problema a ser estudado	26
3.2. Dicionário ou glossário?	26
3.3. Relevância da proposta para os estudos da tradução em geral e melhor conhecimento de um de seus aspectos em particular	27
<b>4. OBJETIVOS</b>	29
4.1. Uso de <i>corpora</i> computadorizados em pesquisas lingüísticas	29
4.2. Importância da análise de <i>corpora</i> para o tradutor	31
4.2.1. Caso da Prosódia Semântica	31
4.2.2. Consulta do <i>corpus</i> específico do tradutor	35
<b>5. METODOLOGIA</b>	37
5.1. Seleção do <i>corpus</i> de análise	38
5.2. <i>Corpus</i> analisado	39
5.3. Preparo dos arquivos	41
5.4. Aplicação do concordanciador WordSmith Tools	42
5.4.1. Listas de palavras	42
5.4.2. Candidatos a termo – princípio geral	43
5.4.3. Candidatos a termo – escolha da autora	46
5.5. Extração de verbetes a partir do <i>corpus</i> de análise	56
5.6. Formato	59
5.7. Sinonímia	60
5.7.1. Apresentação de sinônimos no glossário técnico	65
5.8. Ordem das entradas	66
5.9. Critérios de qualidade	66
5.10. Regionalismos	67
5.11. Possíveis diferenças	71
<b>6. RESULTADOS</b>	73
<b>7. CONSIDERAÇÕES FINAIS</b>	76
<b>8. REFERÊNCIAS</b>	79

## LISTA DE ILUSTRAÇÕES

Número	Descrição	Página
1	Ocorrências do IL "sistema" no <i>corpus</i> em LP	43
2	Ocorrências do IL "system" no <i>corpus</i> em LI	44
3	Ocorrências do IL "ar" no <i>corpus</i> em LP	52
4	Ocorrências do IL "air" no <i>corpus</i> em LI	53
5	Ocorrências do IL "alavanca" no <i>corpus</i> em LP	55
6	Ocorrências do IL "lever" no <i>corpus</i> em LI	56
7	Ocorrências do IL "air conditioning" no <i>corpus</i> em LI	59
8	Ocorrências do IL "air conditioner" no <i>corpus</i> em LI	60
9	Ocorrências do IL "ar condicionado" no <i>corpus</i> em LP	61
10	Ocorrências do IL "condicionador de ar" no <i>corpus</i> em LP	62
11	Ocorrências do IL "elétr*" no <i>corpus</i> em LP	66
12	Ocorrências do IL "electric" no <i>corpus</i> em LI	67
13	Ocorrências do IL "power" no <i>corpus</i> em LI	68

## LISTA DE TABELAS

Número	Descrição	Página
1	Detalhamento do <i>corpus</i>	39
2	100 IL mais freqüentes nos 5 documentos escritos em LP	46
3	100 IL portadores de sentido mais freqüentes nos 5 documentos escritos em LP	47
4	100 IL mais freqüentes nos 9 documentos escritos em LI	48
5	100 IL portadores de sentido mais freqüentes nos 9 documentos escritos em LI	49
6	Extração de verbetes no glossário técnico-automobilístico bilíngüe português ? inglês	57
7	Extração de verbetes no glossário técnico-automobilístico bilíngüe inglês ? português	57
8	Exemplo de verbete no glossário técnico-automobilístico português ? inglês	63
9	Exemplo de verbete no glossário técnico-automobilístico inglês ? português	63
10	Regionalismos	68

## **LISTA DE ABREVIATURAS**

IL: Item lexical

LI: Língua Inglesa

LP: Língua Portuguesa

MT: Manual de Terminologia da Direção de Terminologia e Normalização do  
Departamento de Tradução do Governo Canadense

PGET: Programa de Pós-Graduação em Estudos da Tradução

WST: WordSmith Tools

## 1. INTRODUÇÃO

### 1.1. Lingüística de *Corpus*Erro! Indicador não definido.

O papel do computador na ciência moderna já é bem conhecido. Mas no estudo da linguagem, os computadores têm oferecido novas perspectivas. Eles nos permitem ver fenômenos antes despercebidos devido às limitações da nossa cognição. O impacto do uso de *corpora* computadorizados na lingüística tem sido associado ao uso do telescópio na astronomia ou ao microscópio na biologia. Mas o estudo da linguagem não teve o seu telescópio ou microscópio; o computador é a ferramenta analítica que torna possível, pela primeira vez, um verdadeiro estudo empírico da linguagem<sup>1</sup>. Ou seja, a partir dos estudos lingüísticos com base em *corpora* computadorizados, o uso e a frequência de palavras e expressões em uma determinada língua são comprovados através de sua ocorrência, tanto individualmente como em conjunto com outras palavras (coocorrência).

Em relação à Lingüística de *Corpus*, nada se compara à vontade dos pesquisadores que até hoje ainda se dedicam ao estudo de *corpora* não legíveis por máquina. No entanto, através do subsídio de um PC doméstico, é possível analisar milhões de palavras de forma mais rápida e mais confiável do que se fosse feito sem uma ferramenta computacional. A taxa de erro é bastante reduzida no caso de tarefas repetitivas, como contagem de palavras, listagem dos termos mais frequentes em um *corpus*, dentre outras. Assim, utilizando exemplos práticos, a autora pretende confirmar que o computador é atualmente a principal ferramenta de trabalho do lingüista.

---

<sup>1</sup> Ver, a esse respeito, Stubbs (1996) e Hoey (1993).

Barros (2004), defende que a Lingüística de *Corpus* tornou-se fundamental às pesquisas em Terminologia e em outros campos lingüísticos. A autora faz questão de citar as palavras de Biderman (2001), transcritas por Barros (2004):

“... Para os lexicógrafos, sobretudo, foi uma revolução. Máquinas mais potentes e programas mais eficientes criaram para o lexicógrafo essa prancheta inestimável para sua arte de construir dicionários. O *corpus* veio possibilitar um confronto entre a teoria e os dados empíricos da língua. De fato, o *corpus* pode mostrar como funciona uma língua natural em escala reduzida”. [grifos da autora]

Os pilares da Lingüística de *Corpus* são a visão da linguagem como sistema probabilístico e a abordagem empirista, destacando a análise de dados provenientes da observação da linguagem, reunidos sob a forma de um *corpus*.

## 1.2. Terminologia baseada em *corpus*

Conforme o Manual de Terminologia (doravante, MT), a palavra Terminologia denomina a disciplina lingüística dedicada ao estudo científico dos conceitos e dos termos usados nas línguas de especialidade. Esta disciplina lingüística dedica-se ao estudo científico dos conceitos e termos usados nas línguas de especialidade. A língua de especialidade é aquela utilizada para proporcionar uma comunicação sem ambigüidade numa área determinada do conhecimento ou da prática, com base em um vocabulário e em usos lingüísticos específicos desse campo, enquanto que a língua comum é aquela que usamos no dia-a-dia.

Rousseau (1998) enfatiza que a utilização de *corpora* valoriza a função lingüística e comunicacional da Terminologia junto com a função cognitiva. Deve-se

considerar o *corpus* como uma amostra dos discursos especializados que se deseja analisar para extrair materiais terminológicos. A importância da análise de *corpora* deve-se ao fato que o discurso científico e técnico pode ser considerado como o local onde são elaborados os tecnoletos<sup>2</sup>. Portanto, é no discurso especializado que se formam ou aparecem os termos, que se manifesta a coocorrência, a fraseologia e os elementos estilísticos próprios de um tecnoleto. Considerando-se a existência de ferramentas de análise de textos e de levantamento terminológico, assim como de volumes cada vez maiores de material disponível para análise, é importante que o *corpus* seja legível por máquina. A informatização permite tratar *corpora* mais volumosos, explorá-los de forma mais completa e reutilizá-los para atualizações ou até mesmo para outros projetos, sem contar no ganho de produtividade obtido, principalmente ao eliminar os riscos de erros de digitação ou transcrição dos dados. Além disso, como será visto através dos exemplos deste estudo, IL considerados apenas como pertencentes à língua geral podem ser mais facilmente identificados como formadores de termos técnicos. Ou seja, a Terminologia não estuda apenas os ILs criados com um fim específico, como “acumulador” ou “alternador”, mas também a coocorrência de itens lexicais da língua geral, formando novos significados e, neste caso, termos técnicos, como “ar” + “condicionado”, criando o termo “ar condicionado”.

### **1.3. Pesquisa em um *corpus* de conteúdo técnico**

Com base na compilação de um *corpus* de conteúdo técnico, é possível extrair listas dos itens lexicais mais frequentes e, a partir disso, dar início à elaboração de um

---

<sup>2</sup> Tecnoleto – sinônimo de língua de especialidade



dicionário técnico. O público-alvo de um dicionário técnico são os tradutores técnicos, os professores e alunos de cursos instrumentais de idiomas, professores e alunos de disciplinas técnicas e científicas, a indústria relacionada ao tema, bem como os jornalistas da área técnica afim.

Por que é importante fazer uma pesquisa com base em *corpus* para a criação de um glossário? Como poderá ser verificado no capítulo 4.2.1., é possível provar que a intuição do lingüista ou do lexicógrafo não é suficiente para estabelecer critérios de elaboração das listas dos itens lexicais que devem constar de um dicionário ou dos exemplos que devem aparecer na seqüência de cada entrada.

A autora atua há cinco anos como tradutora técnica, dedicando-se principalmente à área automobilística. Além disso, trabalhou durante dezessete anos em empresas nacionais ou estrangeiras que tinham as relações internacionais como parte importante de sua atividade. Durante a sua própria experiência, a autora sempre se confrontou com as mesmas dificuldades enfrentadas pelos falantes não-nativos de uma língua: ser obrigada a desconfiar das equivalências sugeridas pelos dicionários bilíngües, tanto devido à falta de exemplos reais de uso, como, principalmente, porque as equivalências sugeridas nem sempre correspondem ao contexto em que um determinado IL é empregado. Ao entrar em contato com a Lingüística de *Corpus*, a autora viu abrir-se um novo rumo para a prática da tradução. A comprovação da utilização de IL e colocações em seus contextos reais de uso faz com que a qualidade do trabalho do tradutor aumente significativamente. Conduzindo esta pesquisa baseada em Lingüística de *Corpus* e Terminologia, a autora não só está contribuindo para a melhoria da qualidade do seu próprio trabalho, mas também pretende que seu estudo sirva de fonte de consulta para outros tradutores e estudantes da prática da tradução.

Como o glossário proposto nesta pesquisa refere-se à área automobilística, será utilizado um *corpus* de manuais de proprietários de veículos automotores nos idiomas inglês e português. Poderá ser visto, no decorrer desta dissertação, que partindo da simples compilação de textos técnicos, como manuais de utilização de veículos automotores, é possível:

- Iniciar a elaboração de uma lista de verbetes e suas correspondências – neste caso, português e inglês – em um determinado campo do saber;
- Incluir exemplos reais de uso para cada verbete apresentado, tanto no idioma de partida quanto no idioma de chegada;
- Apresentar, quando pertinente, diferenças regionais em relação ao verbete, bem como possível sinonímia.

## 2. REVISÃO DA LITERATURA

Para a elaboração deste documento, foram feitas pesquisas em publicações na área de Lingüística de *Corpus* e Terminologia, dentre as quais destacam-se as obras abaixo. No que concerne à pesquisa na área de Terminologia propriamente dita, já estão disponíveis interessantes publicações nacionais; apesar de existirem em menor número, as publicações existentes no Brasil são recentes e de extrema relevância, principalmente porque fazem um apanhado do histórico e das atuais correntes do estado da arte neste campo do saber, servindo não apenas aos pesquisadores, mas também aos iniciantes da graduação.

Este capítulo tem o objetivo de resumir as principais obras consultadas para a elaboração desta pesquisa. Os autores são mencionados no decorrer do estudo e, afortunadamente, muitas vezes diferentes autores compartilham do mesmo ponto de vista, reforçando o embasamento teórico deste estudo.

### I. Lingüística de *Corpus*: Tony Berber Sardinha (2004).

Esta é principal obra publicada em língua portuguesa sobre Lingüística de *Corpus*, lançada em 2004. Não é possível realizar um estudo baseado em *corpus* sem antes conhecer as bases desta disciplina. Berber Sardinha, Professor da PUC de São Paulo, faz uma compilação sobre o histórico da Lingüística de *Corpus*, apresenta o seu desenvolvimento e atual situação no Brasil, dando diretrizes para os estudos baseados nesta área do saber.

“Lingüística de *Corpus*” é atualmente a obra mais consultada no País no que se refere a tamanho do *corpus*, como pode ser observado em vários estudos publicados

recentemente. Em relação ao programa WordSmith Tools, o autor apresenta o passo a passo para que qualquer estudante ou pesquisador possa entender como funciona o programa, podendo extrair listas de palavras e concordâncias de um *corpus*. A vantagem em relação à seção “Ajuda” do programa é que as explicações de Berber Sardinha estão apresentadas em Língua Portuguesa.

Berber Sardinha também dá continuidade e sustenta a teoria da Prosódia Semântica, defendida por Allan Partington, dando exemplos contrastivos em língua portuguesa e em língua inglesa. Por último, mas não menos importante, Berber Sardinha apresenta, em um dos anexos de sua obra, uma lista com as três mil palavras mais frequentes do Banco de Português, que pode ser útil para servir de *corpus* de controle da língua geral para estudos baseados em *corpus* de domínio técnico.

## II. Curso Básico de Terminologia, de Lídia Almeida Barros (2004).

Na opinião da autora, é a obra mais esclarecedora e bem elaborada nesta área de abrangência. Lídia Almeida Barros, professora da UNESP de São José do Rio Preto, não só descreve a identidade científica da Terminologia, como os tipos de obras lexicográficas e terminográficas, componentes estruturais dos repertórios, metodologia do trabalho terminológico e aspectos informáticos relacionados. Barros encaminha o pesquisador na metodologia de trabalho de pesquisa terminológica. O capítulo 6 é particularmente importante no que se refere à importância do emprego da Lingüística de *Corpus* no estudo da Terminologia. Apesar de outros autores citarem esta importância, Barros dedicadamente exemplifica, de forma esclarecedora, como os estudos atuais relacionados à Terminologia necessariamente devem passar pela análise informática das bases de dados pesquisadas.

Segundo a autora, os dois primeiros itens constituem a bibliografia mínima para um pesquisador na área de Terminologia baseada em *corpus*, ou até mesmo para um estudante de graduação interessado na disciplina de Terminologia. A autora acredita que, atualmente, a Lingüística de *Corpus* constitui um estudo preliminar ao estudo da Terminologia.

III. Introdução à Terminologia, de Maria da Graça Krieger e Maria José Bocorny Finatto (2004).

Maria da Graça Krieger e Maria José Bocorny Finatto, professoras da Universidade Federal do Rio Grande do Sul, destacam a fundamentação da Terminologia e traçam um panorama da recente produção científica sobre o tema. O ano da primeira publicação desta obra coincide com a mencionada no item II, mas as duas se diferem devido à primeira dar ênfase às aplicações práticas da Terminologia e a segunda ressaltar o histórico da Terminologia em âmbito internacional e nacional, contribuindo, também, para validar a Terminologia como disciplina autônoma.

IV. Manual de Terminologia, de Silvia Pavel e Diane Nolet (2002).

Documento on-line disponibilizado pelo Office Québécois de La Langue Française, do Canadá, que apresenta os princípios da pesquisa terminológica, a metodologia e as ferramentas de trabalho disponíveis on-line e off-line, tanto para utilização livre quanto mediante pagamento. O complemento ideal deste manual é a análise das aplicações práticas da Terminologia no próprio site do Grand Dictionnaire Terminologique, também do Office Québécois de la Langue Française.

V. Terminologia: questões teóricas, métodos e projectos, coordenação de Maria Helena Mateus e Margarita Correia (1998).

Compilação de artigos nos idiomas português, espanhol e francês, organizada por Maria Helena Mateus e Margarita Correia, ambas da Universidade de Lisboa. Destaque para os artigos do tradutor literário e principal promotor da Terminologia em Cuba, Rodolfo Alpízar Castillo, e de Louis-Jean Rousseau, Chefe do Departamento de Trabalhos Terminológicos do Office Québécois de la Langue Française que, na opinião da autora, são pesquisadores da atualidade que se destacam na aplicação prática da Terminologia.

VI. Patterns and Meanings, de Alan Partington (1998).

Partington aborda estudos de *corpora* e concordância, colocações, fraseologia. Para o estudo aqui apresentado, a principal contribuição desta obra deve-se à aplicação da prosódia semântica na tradução e na elaboração de dicionários, apresentada no capítulo 4, “Connotation and semantic prosody”, como poderá ser visto em 4.2.1., abaixo.

VII. As Ciências do Léxico, organizado por Ana Maria Pinto Pires de Oliveira e Aparecida Negri Isquerdo (2001).

Compilação de artigos que tratam de Lexicologia, Lexicografia e Terminologia, organizada pelas professoras da Universidade Federal de Mato Grosso do Sul.

Destaque para o artigo de Claudia Maria Xatara, sobre Dicionários Bilíngües e o Problema da Tradução.

- VIII. Computational Approaches to the Lexicon, organizado por B.T.S. Atkins e A. Zampolli (1994), com ênfase para o artigo de Makoto Nagao, professor da Universidade de Kioto: A Methodology for the Construction of a Terminology Dictionary.

Makoto Nagao dedica-se à pesquisa em Processamento de Linguagem Natural, análise de *corpus* monolíngüe e bilíngüe e construção de dicionários automatizados, entre outras. Neste artigo, Nagao demonstra passo a passo como organizar as informações para a elaboração de um dicionário terminológico (não baseado em *corpus*).

- IX. Practical Lexicography, de Bo Svensén (1993).

Obra completa para o fazer dicionarístico. O sueco Svensén retrata desde os tipos de dicionários e seus usuários-alvo como também aspectos de construção, colocações, definições, informações enciclopédicas, etimologia, microestrutura.

### 3. JUSTIFICATIVA

#### 3.1. Problema a ser estudado

Apesar dos estudos lingüísticos com base em *corpora* computadorizados estarem razoavelmente avançados, de maneira especial na Europa, ainda não há aplicações comerciais para utilização em larga escala com foco em Terminologia, principalmente no Brasil. Na introdução do dicionário de língua inglesa (doravante, LI) Collins COBUILD (*Collins Birmingham University International Language Database*), está indicado “... *technical or scientific textbooks, manuals, directories, and so on are not included in the corpus*” (livros-texto, manuais, diretórios e outros documentos técnico-científicos não estão incluídos no *corpus* [de análise] – tradução da autora).

Mesmo no caso dos dicionários organizados com base em *corpus*, a listagem dos verbetes é escolhida com base em sua frequência na língua. Como os termos técnicos são menos frequentes do ponto de vista de um *corpus* genérico, um dicionário de termos técnicos precisa ser elaborado com base na análise de um *corpus* específico, formado por textos técnicos, para que a frequência e o emprego destes termos possam ser analisados.

O profissional de tradução, notadamente aquele especializado em traduções técnicas, dispõe de poucas ferramentas para auxiliá-lo na decisão sobre o melhor termo a ser utilizado para substituir um IL de âmbito técnico. Os poucos dicionários técnicos bilíngües existentes, principalmente destinados a falantes do português do Brasil, podem apresentar problemas de confiabilidade, visto que não são baseados no uso das expressões em seu contexto de utilização e nem em seus padrões de coocorrência.



Os termos técnicos também diferem, dependendo do âmbito de aplicação (engenharia elétrica, engenharia mecânica, engenharia naval, etc.). Um exemplo é o substantivo *chicote*. Em português, um dos sentidos para esta palavra está relacionado a *cabo elétrico*, muito utilizado na indústria automobilística no que se refere à conexão de acessórios elétricos equipados em veículos automotores (auto-rádio, alto-falantes, etc.). Procurando por este termo em um dicionário português ? francês, encontramos como significado técnico a expressão *bout de câble*. Entretanto, após uma investigação mais detalhada, apreende-se que *bout de câble* quer dizer “extremidade de um cabo”, sendo este termo empregado apenas na área náutica<sup>3</sup>.

Outra questão a ser considerada ao trabalhar com Terminologia é a variação sócio-cultural. Muitas vezes os operários e técnicos de nível médio utilizam a sua própria “terminologia” no seu cotidiano de trabalho. A autora, durante a prática da tradução, já vivenciou circunstâncias inusitadas. Por exemplo, o termo em português “boqueteiro” é utilizado habitualmente pelos colaboradores de uma empresa no Brasil. Ao fazer uma pesquisa na Internet com este termo, especificando “somente páginas em português”, encontramos pouco mais de 300 ocorrências, em sua grande maioria relacionadas à pornografia. Entretanto, este adjetivo é uma adaptação do substantivo “boquette” (gíria), relacionado a “sala”, “escritório”. Em uma concessionária de veículos, o “boqueteiro” é o responsável pela pequena sala onde são armazenadas as peças de reposição para utilização na oficina. Segundo Albuquerque e Cerceau (1996),

“A língua é viva e mutável e se confunde com as necessidades de seus falantes e usuários. Na empresa, as pessoas que sempre trabalharam com determinados equipamentos siderúrgicos começam a se familiarizar com eles e acabam criando associações entre tais equipamentos e coisas não relacionadas à siderurgia. Criam uma linguagem própria, ou seja, atuam sobre a língua, dando vida para ela. Esse é o caso do defletor iniciador de

---

<sup>3</sup> Em francês, *chicote* traduz-se por *faisceau de fils*, conforme o Grand Dictionnaire Terminologique.

bobina que, dentro da CSN, transformou-se em bico de pato, ou do rolo controlador de tensão, conhecido como rolo dançarino, e do equipamento chamado *belt wrapper*, cuja tradução no dicionário Taylor significa cinta de bobinar tiras e na CSN é conhecido como vaca. Para alguns técnicos, não há ainda uma tradução adequada para nossa língua desse termo”.

As obras terminográficas contribuem para a defesa da identidade de um idioma e do seu povo. No âmbito industrial, os profissionais com menos educação formal precisam saber do que estão falando quando se referem a um componente. Não precisamos dizer *powertrain* se podemos dizer “grupo motopropulsor”, até porque é mais fácil pronunciar uma palavra do seu próprio idioma do que termos de um idioma estrangeiro.

### **3.2. Dicionário ou Glossário?**

Este assunto causa muita polêmica entre terminógrafos e terminólogos, mas, *a priori*, define-se glossário como um repertório de unidades lexicais de uma especialidade, sem pretensão de exaustividade. Já o dicionário terminológico oferece conceitos e se caracteriza por uma cobertura exaustiva de itens lexicais. Esta metodologia objetiva a apresentação de listas de termos em dois idiomas com exemplos reais de utilização em cada idioma retratado. A língua de partida é o português do Brasil.

Mesmo que o lexicógrafo ou o terminógrafo seja muito experiente, não basta apenas a sua intuição para que ele saiba quais termos devem fazer parte do glossário e quais não. Assim como o fazem os dicionaristas que utilizam a análise de *corpora* legível por máquina para a elaboração dos grandes dicionários, a análise de *corpora* também deve ser a base para a elaboração de glossários técnicos. Um passo além seria

incluir no glossário apenas exemplos extraídos do *corpus* de análise, de forma a enriquecer o material de consulta para o profissional da tradução. Cabe ressaltar que, no caso aqui apresentado, a autora considera dispensável a inclusão de definições, visto que o produto desta metodologia é um glossário bilíngüe.

### **3.3. Relevância da proposta para os estudos da tradução em geral e melhor conhecimento de um de seus aspectos em particular**

A análise lingüística de *corpora* bilíngües legíveis por máquina abre caminho para diversos tipos de pesquisas, tendo como objetivos principais :

- Facilitar a compreensão sobre as línguas comparadas, cujas particularidades podem passar despercebidas caso sejam estudadas individualmente;
- Aumentar o nosso conhecimento sobre diferenças culturais, tipológicas e específicas de cada língua, assim como características comuns;
- Destacar as diferenças entre textos originais e traduções, e entre textos produzidos por falantes nativos e não-nativos;
- Aplicações práticas como, por exemplo, lexicografia, ensino de línguas e tradução.

Nota-se que há poucos estudos voltados à prática da tradução, principalmente especializados em tradução técnica ou não relacionados à tradução literária. Já existe um número razoável de obras relacionadas à Terminologia no Brasil, mas poucas

mostram na prática os passos seguidos e as dificuldades encontradas por um profissional da tradução durante a criação do seu próprio banco de dados terminológico.

O profissional de tradução especializada, que presta serviços para várias empresas simultaneamente, tem em geral uma grande carga de trabalho e prazos muito curtos para executar o seu trabalho. Desta forma, resta pouco tempo para que ele possa fazer buscas muito longas em materiais previamente traduzidos ou até mesmo para criar o seu próprio glossário de termos técnicos, de acordo com a sua área de atuação.

Por outro lado, o tradutor pode estar mais envolvido na produção formal de dicionários do que se pode imaginar. De acordo com Barros (2004),

“A produção de dicionários por tradutores ao longo da história da humanidade é enorme, vai dos glossários médicos da Antigüidade aos moderníssimos dicionários informatizados e aos bancos de dados com acesso via Internet. A cooperação entre tradutores e terminólogos, ou mais particularmente o trabalho dos tradutores como terminólogos, pode ser testemunhado por inúmeras obras terminográficas bilíngües ou multilíngües, elaboradas em épocas diferentes, tanto no Ocidente quanto no Oriente”.

## 4. OBJETIVOS

Esta pesquisa visa a apresentar uma metodologia para a elaboração de um glossário bilíngüe com base em um *corpus* do domínio técnico-automobilístico.

### 4.1. Uso de *corpora* computadorizados em pesquisas lingüísticas

Entende-se por *corpora* computadorizados (ou *corpora* legíveis por máquina) aqueles cujo formato é passível de leitura por máquina. A análise de *corpora* computadorizados é feita através de programas computacionais específicos que executam as operações de etiquetagem<sup>4</sup>, listagem de palavras e concordância, entre outras.

Segundo Berber Sardinha (2004), para que a análise de *corpora* computadorizados tenha validade, algumas condições devem ser preenchidas:

- a. Os textos que formam o *corpus* devem ser provenientes de linguagem natural e não produzidos com o propósito de serem analisados em investigações lingüísticas.
  - Neste estudo, serão utilizados manuais de orientação para proprietários de veículos automotores, criados pelos seus fabricantes.

---

<sup>4</sup> Segundo Berber Sardinha (2004), a etiquetagem do corpus consiste na inserção de informações referentes a cada unidade do texto (morfológica, sintática, semântica, discursiva).

- b. Os textos devem ser produzidos por falantes nativos, a não ser que o estudo tenha como objetivo a análise de *corpora* de aprendizes, mais conhecidos como *learner corpora*. No caso de análises contrastivas, utilizam-se originais e traduções.
- Para este estudo, serão utilizados textos originais, em inglês, provenientes dos Estados Unidos e da Inglaterra. Para uma análise contrastiva, será utilizado um texto original, em português, produzido por uma montadora de veículos brasileira<sup>5</sup>. Este texto será comparado com a tradução de um dos manuais produzidos em país de língua inglesa.
- c. O conteúdo deve ser representativo de um idioma ou de uma variedade lingüística.
- Neste estudo, serão analisados textos representativos da variedade textual manuais de proprietários de veículo nos idiomas inglês (dos Estados Unidos e da Inglaterra) e português (do Brasil).

Dois tipos principais de *corpora* podem ser utilizados em análise contrastiva: (i) aqueles formados por textos originais produzidos em duas ou mais línguas (*corpora* comparáveis), e (ii) aqueles formados por textos originais e respectivas traduções (*corpora* de tradução).

*Corpora* comparáveis são formados por textos originais de cada língua, selecionados por época de criação, domínio, gênero, público-alvo, etc. Eles devem

---

<sup>5</sup> As montadoras de veículos brasileiras são, em sua maioria, subsidiárias de empresas com sede na América do Norte, Europa e Extremo Oriente. Por este motivo, os veículos (os seus produtos) são nacionalizados (ou adaptados para produção e comercialização local). Assim, os manuais de utilização destinados aos proprietários destes produtos (os veículos) são traduções adaptadas dos originais, escritos no idioma de origem das matrizes ou, por conveniência no caso de empresas que comercializam os seus produtos em vários países, em inglês.

representar a língua natural de cada idioma analisado. Neste estudo, serão utilizados textos originais em língua inglesa e portuguesa, bem como três textos traduzidos de um dos originais do inglês para o português.

Alguns problemas podem ser encontrados quando se utilizam *corpora* de tradução para análises contrastivas. Dependendo dos idiomas de origem e de destino e do domínio a que se referem (técnico, literatura, etc.), pode haver uma certa dificuldade na obtenção de originais e traduções que sejam suficientemente representativos de cada língua. No caso do português do Brasil, principalmente em relação ao domínio técnico-científico, pode-se afirmar que originais em inglês e traduções para o português são encontrados com mais frequência do que o inverso.

## **4.2. Importância da análise de *corpora* para o tradutor**

Nesta subseção serão apresentados dois aspectos da pesquisa em tradução com base em *corpora* computadorizados: o caso da prosódia semântica e a consulta do *corpus* específico do tradutor.

### **4.2.1. Caso da Prosódia Semântica**

Normalmente um tradutor faz freqüentes pesquisas na Internet para encontrar novas palavras e expressões ou investigar o seu uso, sejam elas técnicas, de gíria, inovações de um determinado idioma, etc. Encarando a Lingüística de *Corpus* de uma forma simplista, seria possível afirmar que estas pesquisas na Internet já representam um primeiro passo na pesquisa com base em *corpora* computadorizados, visto que, ao pesquisar textos produzidos por falantes nativos na Internet, já estamos buscando o real

emprego de palavras e expressões. Este tipo de pesquisa simples já difere do que é proposto pelos dicionários em geral, que são normalmente baseados na gramática tradicional e combinados a uma certa liberdade em relação à lexicografia, como se pode verificar através do prefácio do Novo Dicionário Aurélio da Língua Portuguesa (1986):

“Quando se pensa em abonar palavras para um léxico, observa-se o mundo dos vocábulos e, sobretudo, de significados, vivos de um século, de séculos, e esquecidos dos lexicógrafos, que, ao longo do tempo (com exceções, é claro), foram fazendo pouco mais do que mutuamente se copiarem (o que às vezes é inevitável), com, aqui e ali, algum retoque na cópia, e mais umas tantas palavras ou acepções, que estão entrando pelos olhos.”

“... e lutam, igualmente, os dicionaristas, redefinindo-as [as palavras], acrescentando-lhes significados, ou introduzindo-as no léxico, após enfrentar a tarefa, tantas vezes penosa, de captar-lhes a essência, desentranhar-lhes o sentido, infundir alma num corpo”.

Robert Allen (1986), em seu artigo *A Concise History of the COD (Concise Oxford Dictionary)*, relata como os dicionários eram elaborados no início do século XX:

“In one important respect the COD was of course very different from the OED [Oxford English Dictionary]: it was supported by invented examples and not by quotations from real sources; this would have required attribution for which there was no space. Invented examples afford a dictionary compiler the best opportunity of a creative fling, and the Fowlers’ examples are in many ways exceptionally distinctive...”

No mesmo artigo, Allen relata uma declaração feita pelo dicionarista e editor Henry Fowler no início do século passado:

“Now that we have got beyond the OED, we go rather slower, produce rather shorter articles, & have always the consciousness, with the words that really matter, that a third or so of the senses are likely to be escaping us.” [Grifos meus]



Mais adiante, Allen relata que os dicionaristas do início do século XX chegavam a se desentender sobre as próprias regras de organização dos dicionários e escolha dos vocábulos a serem incluídos. Utilizando expressões como “eu acho”, “eu sugiro”, os editores responsáveis seguiam com a determinação sobre os verbetes que eram ou não eram importantes a ponto de serem incluídos em uma nova edição do dicionário. Mais uma vez, Allen transcreve uma declaração feita em uma carta de 28 de setembro de 1928 de autoria de Fowler:

“I should say, let us cut out the word & have no more fuss about it.”

A partir dos anos 70, a Oxford University Press decidiu rever todo o processo de organização dos seus dicionários. Contratou um editor que, além de tradutor técnico, também era pesquisador e astrofísico, levando uma nova forma de pensar ao COD: a visão científica que admite que dados devem ser sondados apenas se verificáveis e quantificáveis, e que considera fundamental a consistência na abordagem de todos os elementos de informação. Em 1980, este editor, J.B. Sykes, declarou em uma entrevista:

“I think I have the kind of mind that likes the precision and deduction of the scientific method, rather than the imprecise and subjective approach that you find in the arts” [Grifos meus]

Com relação à atual proposta de dicionários, glossários e tesouros<sup>6</sup>, Berber Sardinha (2004) e Partington (1998) descrevem em detalhe o problema não revelado a

---

<sup>6</sup> Este termo já está sendo empregado em lugar de "tesouros", conforme exibido nas páginas do Projeto Linguatca na Internet: [www.linguatca.pt](http://www.linguatca.pt).

falantes não-nativos de um idioma e, principalmente, aos “tradutores ingênuos”<sup>7</sup>, quando consultam estas ferramentas de consulta. Trata-se da prosódia semântica.

Conforme Berber Sardinha (2000), a prosódia semântica refere-se à associação entre itens lexicais e conotação (negativa, positiva ou neutra) de campos semânticos. O nome deve-se ao fato de certas palavras prepararem o ouvinte ou o leitor para o conteúdo semântico que está por vir, da mesma maneira que a prosódia na fala indica para o interlocutor que tipos de sons estão por vir a seguir. Por exemplo, ‘cause’ tem uma prosódia semântica negativa, pois se associa a palavras desfavoráveis como ‘problem(s)’, ‘damage’, ‘death(s)’, ‘disease’, ‘concern’ e ‘cancer’. Já ‘provide’ possui uma prosódia semântica positiva ou neutra, já que se associa a palavras deste tipo, tais como ‘assistance’, ‘care’, ‘jobs’, ‘opportunities’ e ‘training’.

Berber Sardinha (2004) cita relato de Partington (1998):

“Por exemplo, o adjetivo inglês *impressive* (impressionante) coloca-se com itens como *achievement*, *talent* e *dignity*. Por isso, parece exibir uma prosódia semântica positiva. Já *rife* (abundante, predominante) tende a coocorrer com palavras como *crime*, *misery* e *diseases*, daí a ser considerado portador de uma prosódia semântica negativa”.

Outro caso bastante interessante relatado por Berber Sardinha toma novamente como exemplo a palavra *impressionante*. Em inglês, conforme relatado acima, a palavra *impressive* é portadora de prosódia semântica positiva. Por outro lado, em italiano, esta palavra é portadora de prosódia semântica negativa, pois está normalmente inserida em contextos referentes a atentados, aumento de preços. Isto é o que comumente se caracteriza por falsos cognatos ou “falsos amigos” (*false friends*).

---

<sup>7</sup> Ver Stella Tagnin (2001).

Comentando relatório elaborado por Harvey & Yuill (1994), Partington (1998) revela que a busca por sinônimos representa dez por cento das consultas feitas a dicionários quando tradutores e aprendizes (ou falantes não-nativos de uma língua) estão envolvidos na redação de um texto. Como a maioria dos dicionários, glossários e tesouros não informam a aplicação correta de cada sinônimo, verificou-se que, em mais de 36% das buscas por sinônimos realizadas por aprendizes, o verbete apresentado pelo dicionário não lhes forneceu a informação de que precisavam. Desta forma, comparar o comportamento de determinadas palavras ou expressões através da pesquisa em *corpora* pode complementar a pesquisa realizada em dicionários, glossários e tesouros.

#### 4.2.2. Consulta do *Corpus* Específico do Tradutor

Tradutores com uma certa experiência em tradução técnica normalmente produzem uma quantidade razoável de textos traduzidos e, portanto, também possuem uma quantidade proporcionalmente considerável de originais (tanto no seu próprio idioma como em idiomas estrangeiros).

Segundo Mona Baker (1995), compilação de terminologia baseada em *corpus* é hoje mais uma regra do que uma exceção. Este método é particularmente útil para o desenvolvimento de um glossário técnico de uso pessoal do tradutor.

Aproveitamos para recordar as considerações de Baker (1995) em relação aos seis critérios de seleção por tipos de *corpora*:

- I. Linguagem genérica ou de domínio restrito (técnico, por exemplo);
- II. Escrito ou falado;

- III. Sincrônico (relativo a uma mesma época a ser determinada) ou diacrônico (relativo a épocas distintas);
- IV. Origens (escritores / falantes) e gêneros (mídia impressa ou eletrônica, literatura, audiências de tribunais);
- V. Limites geográficos (inglês americano x inglês britânico, ou português do Brasil x português europeu);
- VI. Monolíngüe, bilíngüe ou multilíngüe.

No estudo aqui apresentado, os itens 1, 4, 5 e 6 merecem destaque. Como poderá ser visto adiante, até mesmo termos técnicos podem ser diferentes em países que falam um mesmo idioma, mesmo nos casos de áreas amplamente abordadas mundialmente por falantes tanto nativos quanto não-nativos. A área automobilística é um bom exemplo, pois está freqüentemente presente na mídia impressa e eletrônica de todo o mundo, sendo inclusive alvo de reportagens, programas de televisão e seções exclusivas em vários meios de comunicação em massa.

De acordo com Baker, normalmente se pensa que qualquer estudo envolvendo tradução deve contrastar *corpora* de dois ou mais idiomas, não podendo, portanto, ser monolíngüe. Entretanto, podem ser contrastadas variantes lingüísticas de diferentes países, como inglês americano e inglês britânico.

## 5. METODOLOGIA

A fim de apresentar como as ferramentas computacionais podem auxiliar o tradutor a organizar o seu próprio glossário técnico, o referencial teórico desta pesquisa é constituído de exemplos de utilização do software WordSmith Tools (doravante, WST).

Reproduzindo Rousseau (1998), a generalização do uso de tecnologias da informação e comunicação acarretou profundas mutações na organização do trabalho terminológico. Uma das principais características da nova organização do trabalho que se delineia é o fato da cadeia de produção da terminologia ser amplamente informatizada em todas as etapas do trabalho, seja na fase documentária, de tratamento ou de divulgação. Além disso, existe a possibilidade de atualizar os trabalhos de forma permanente.

Através da análise de *corpus* por meio de ferramentas computadorizadas, é possível decidir quais IL são relevantes para a inclusão em um glossário de termos técnicos. Sem o auxílio de computadores, esta tarefa seria assaz trabalhosa e poderia conduzir a equívocos, além de custar muito tempo ou o trabalho de vários profissionais trabalhando em um mesmo projeto.

Para o tradutor, a análise do *corpus* também se torna relevante do ponto de vista do contexto em que são empregados os termos técnicos. Através do estudo de um *corpus* especializado na área técnica automobilística, podem ser extraídas valiosas informações sobre as palavras específicas que são encontradas próximas aos itens lexicais em análise. Quando da criação deste glossário técnico, tais exemplos deverão ser transcritos na seqüência de cada verbete, otimizando o emprego dos mesmos pelo

“aprendiz” (falante não-nativo). Assim, quando os verbetes são acompanhados de exemplos originados de textos reais, tornam-se confiáveis por se tratarem de utilização em um contexto autêntico de uso e não exemplos criados para os fins do próprio verbete.

Além disso, se os verbetes escolhidos para serem incluídos em um glossário técnico forem originados da observação direta de uma lista de frequências de um *corpus* específico, estes serão os termos que um usuário ou tradutor mais provavelmente precisará saber.

### 5.1. Seleção do *corpus* de análise

Alguns critérios básicos foram obedecidos para a seleção do *corpus* de análise:

- Origem: Uma das formas de se controlar a qualidade dos textos é utilizar apenas publicações, visto que textos publicados normalmente correspondem a um certo padrão de aceitabilidade no idioma em que foram escritos.

Os textos que serviram a esta análise referem-se a transcrições legíveis por máquina de textos originais publicados por montadoras de veículos para os seus clientes compradores desses mesmos veículos.

- Direitos Autorais: Paralelamente à escolha do tipo de texto, deve-se ter em mente que, caso a análise tenha como objetivo a publicação de um artigo técnico-científico ou a elaboração de um glossário técnico, deve ser solicitada aos respectivos autores a permissão para a reprodução do material consultado. A maior parte dos autores especialistas no assunto afirma que é mais fácil conseguir

permissão para utilização de trechos de documentos do que textos completos. Por outro lado, no caso de estudos individuais / particulares, ou seja, um tradutor analisando um *corpus* de domínio específico para o seu próprio uso, não deve haver uma preocupação com direitos autorais.

- Organização: Os textos devem ser organizados e salvos em formato legível por máquina. Os formatos legíveis pela maior parte dos concordanciadores<sup>8</sup> são .txt (arquivos de texto puro sem formatação) ou .html (linguagem de marcação padrão usada em documentos na Internet).

## 5.2. *Corpus* analisado

Foram analisados catorze documentos de teor técnico (manuais e folhetos), totalizando 381.359 palavras. O *subcorpus* em LI contém nove documentos (285.297 palavras) e o *subcorpus* em LP totaliza cinco documentos (96.062 palavras). Os documentos foram cedidos para análise particular pelas empresas que os produziram em meio digital (CD).

1. Ford Escort – Owner's Guide – 2001 (Estados Unidos);
2. Ford EcoSport – Owner's Guide – 2004 (Estados Unidos);
3. Ford EcoSport – Manual do Proprietário – 2004 (Brasil);
4. Nissan Altima – Owner's Manual – 2002 (Estados Unidos);
5. Nissan Pathfinder – Owner's Manual – 2005 (Estados Unidos);
6. Nissan Sentra – Owner's Manual – 2004 (Estados Unidos);
7. Renault Clio – The Driver's Handbook – 2001 (Inglaterra);

---

<sup>8</sup> Concordanciador: Conforme o MT, trata-se de um programa que identifica e enumera as coocorrências dos termos nos textos selecionados durante a pesquisa temática.

8. Renault Clio – Manual de Utilização – 2002 (Brasil);
9. Renault Scénic – The Driver's Handbook – 1999 (Inglaterra);
10. Renault Scénic – Manual de Utilização – 2003 (Brasil);
11. Toyota Corolla – Pocket Reference Guide – 2005 (Estados Unidos);
12. Troller – Manual do Proprietário – 2004 (Brasil);
13. Volkswagen (Understanding your model year 2000 Volkswagen Features; Multi-Function Steering Wheel; Automatic Transmission Operation; Understanding your Vehicle's Ventilation System; Understanding How to Minimize Wind Noise) – 2000 (Estados Unidos);
14. Volkswagen SP2 (Freios – verificação e regulagem) – 2005 (Brasil).

Inglês			Português		
Fabricante - Modelo	Autoria	Disposição	Fabricante - Modelo	Autoria	Disposição
Ford - Ecosport	Original	Paralelo	Ford - Ecosport	Tradução	Paralelo
Ford - Escort	Original	-	-	-	-
Nissan - Altima	Original	-	-	-	-
Nissan - Pathfinder	Original	-	-	-	-
Nissan - Sentra	Original	-	-	-	-
Renault - Clio	Original	Comparável	Renault - Clio	Tradução	Comparável
Renault - Scénic	Original	Comparável	Renault - Scénic	Tradução	Comparável
Toyota - Corolla	Original	-	-	-	-
-	-	-	Troller - Troller	Original	-
Volkswagen – Folhetos técnicos (6)	Original	-	Volkswagen – Folheto técnico (1)	Tradução	-

Tabela nº 1: Detalhamento do *corpus*

Com base na classificação realizada por Berber Sardinha (2004), trata-se de um *corpus* bilíngüe de tamanho médio.

### 5.3. Preparo dos arquivos



Os arquivos foram analisados através de um programa de concordância (concordanciador). Ao utilizar um concordanciador para pesquisa terminológica, o programa pode contar e enumerar todos os casos em que aparece um determinado termo, mostrando-o junto com as palavras que o antecedem e o sucedem, de acordo com os limites estabelecidos pelo usuário (como, por exemplo, 10 palavras antes e 10 palavras depois do termo). A partir da lista de ocorrências apresentada pelo programa em relação a um determinado termo, o terminógrafo identifica os ILs mais frequentes.

Para que possam ser analisados por um concordanciador, os arquivos devem ser salvos preferencialmente em formato texto. Entretanto, documentos muito extensos ou, principalmente, documentos técnicos que combinam texto e figuras, como é o caso dos manuais, são normalmente salvos em formato .pdf. Como todos os arquivos obtidos estavam gravados em formato Adobe (.pdf), os seguintes cuidados foram tomados:

- I. Cada arquivo foi aberto na versão *freeware*<sup>9</sup> do Adobe Reader 7.0. Na barra de ferramentas deste programa (aba “Arquivo”), há uma opção “Salvar como texto”. Assim, para cada arquivo original em formato .pdf, foi criado um novo arquivo em formato .txt.
- II. Os arquivos originais em formato .pdf foram mantidos para eventual consulta e como cópia de segurança em um diretório específico denominado “Originais em PDF”.
- III. No caso de manuais desdobrados em vários arquivos (às vezes um para cada capítulo), todos os arquivos referentes a um mesmo manual foram copiados e colados para um único arquivo. Conseqüentemente, o número de arquivos em

---

<sup>9</sup> Software distribuído em regime gratuito, mas com algumas restrições.

formato .txt ficou menor que o número de arquivos em formato .pdf, visto que vários originais estavam divididos em diversos arquivos.

- IV. Os arquivos em formato .txt foram salvos com nomes contendo no máximo oito dígitos, visto que o concordanciador WST não reconhece nomes de arquivos a partir do nono caractere. Exemplo: clioport.txt (manual do veículo Clio, língua portuguesa).
- V. Devido ao limite de caracteres para nomear os arquivos em formato .txt, os nomes dos arquivos contêm informações tanto sobre o nome do modelo do veículo quanto uma referência ao idioma em que o manual foi escrito.

#### **5.4. Aplicação do concordanciador WordSmith Tools**

Conforme já citado, algumas ferramentas computacionais auxiliam o analista nos estudos e na prática da tradução. O programa *WordSmith Tools*, de autoria de Mike Scott (1999), obtido através da Internet, foi o software utilizado neste estudo.

##### **5.4.1. Listas de palavras**

O primeiro passo para organizar um glossário a partir de um *corpus* é relacionar todos os ILS contidos nele. Através da lista de palavras apresentada pelo recurso WordList (ou Lista de Palavras Individuais) do programa WST, bem como da análise de suas coocorrências, pode-se explorar um *corpus* de domínio técnico para que sejam tomadas decisões sobre quais palavras devem figurar em um glossário de termos técnicos.

#### 5.4.2. Candidatos a termo – princípio geral

Alguns terminólogos afirmam que uma das maneiras de se decidir quais termos devem fazer parte de um glossário especializado é fazer uma comparação entre a lista de palavras de um *corpus* genérico com um *corpus* de domínio especializado. Em um primeiro momento, são listados os ILs contidos no *corpus* de domínio especializado através do recurso Listas de Palavras do programa. Em seguida, são listados os ILs contidos no *corpus* genérico, através do mesmo procedimento. Procede-se a um cruzamento das duas listas. Os ILs que se repetem em ambas as listas de palavras devem ser excluídos. Conclui-se, assim, que os itens lexicais que não se repetem em ambas as listas de palavras são de domínio técnico e, portanto, devem ser listados no glossário técnico.

Apenas como exemplo, analisando as três mil palavras mais frequentes do Banco de Português listadas por Berber Sardinha (2004), foi encontrado o item “sistema” em 203º lugar, com 52.219 ocorrências. No *subcorpus* analisado em português para esta pesquisa, “sistema” ocupa o 28º lugar (367 ocorrências) e “sistemas” ocupa o 285º lugar (45 ocorrências). Abaixo seguem alguns exemplos do emprego do IL “sistema” (português) / “system” (inglês):

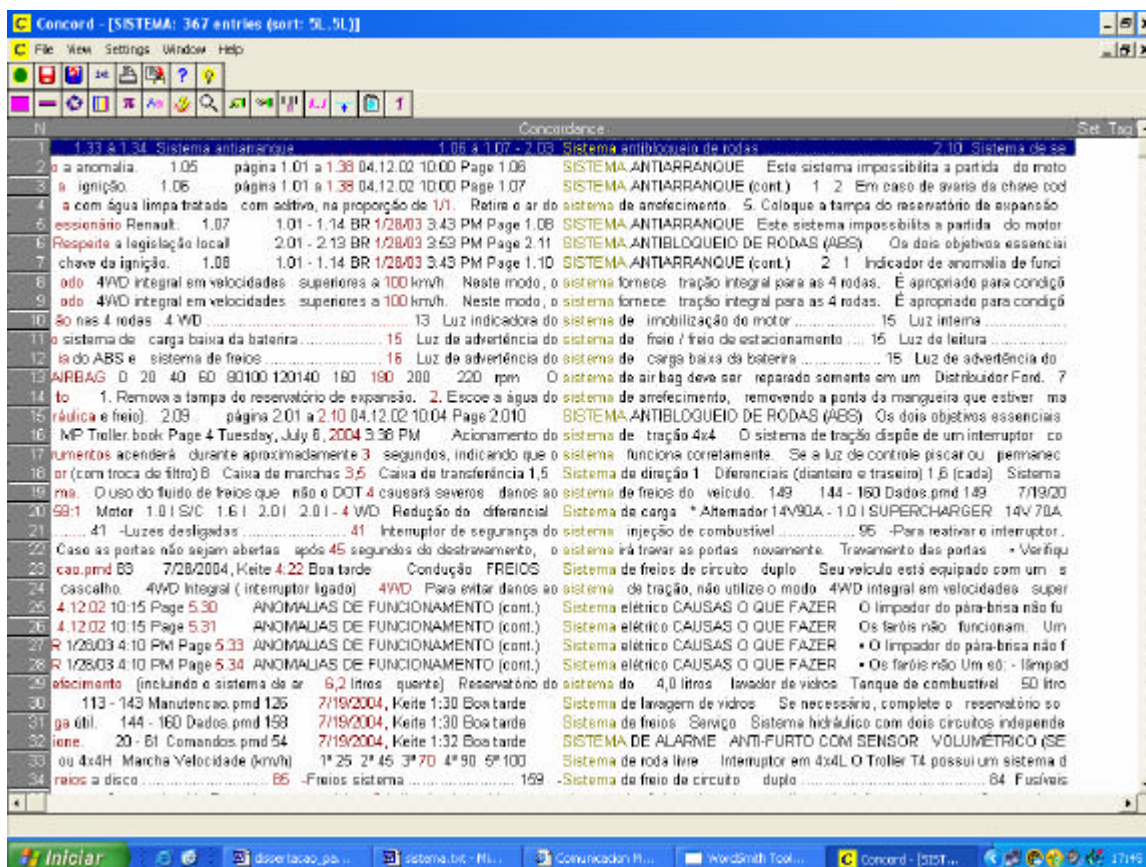


Figura nº 1: Ocorrências do IL “sistema” no *corpus* em língua portuguesa

No *corpus* em português, o IL “sistema” figura em colocações como “sistema antibloqueio de rodas (ABS)”, “sistema elétrico”, “sistema de lavagem de vidros”, “sistema de alarme antifurto com sensor volumétrico”. Portanto, não pode ser descartado da relação de entradas que devem ser listadas no glossário técnico.

Vejam agora como o IL “system” se comporta no *corpus* em LI:

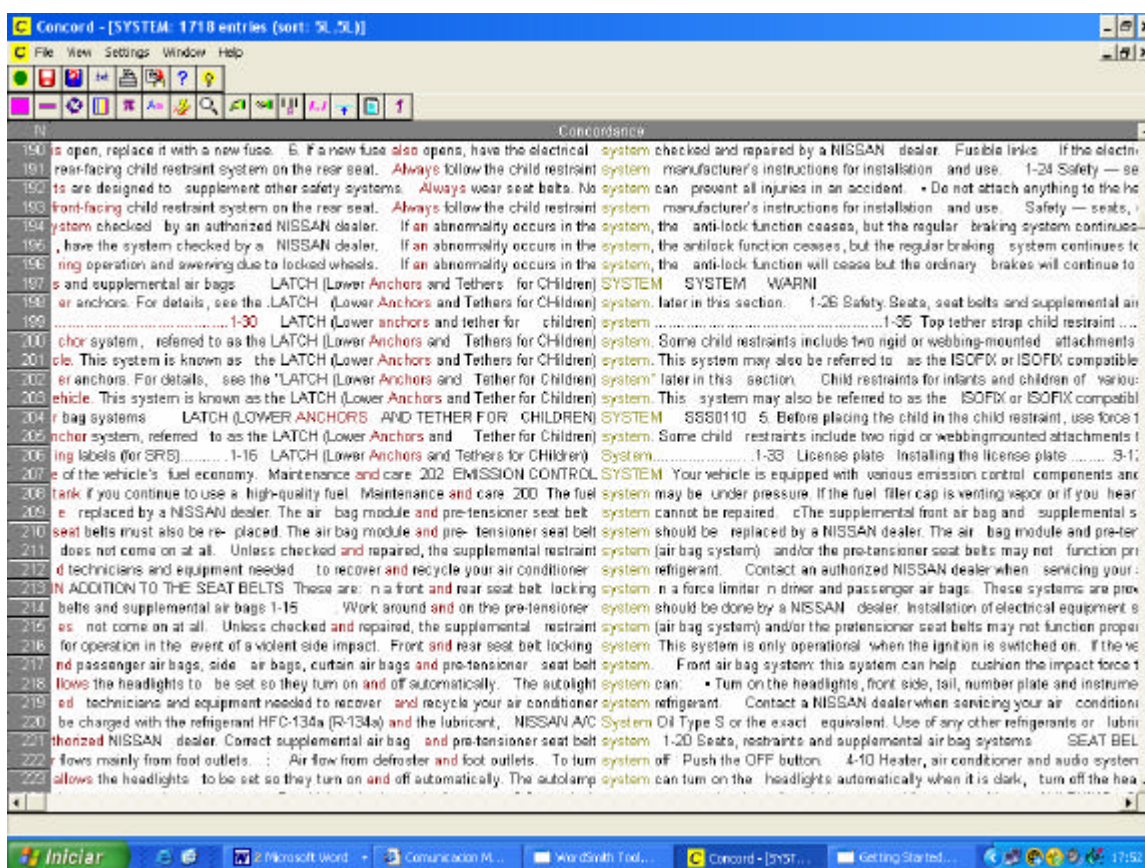


Figura nº 2: Ocorrências do IL “system” no *corpus* em LI.

Observando o *corpus* em LI, “system” figura em colocações como “electrical system”, “child restraint system”, “LATCH (lower anchors and tethers for children) system”, “emission control system”. Assim como em português, o IL system deve ser incluído na lista de candidatos a termo de um glossário técnico da área automobilística, tendo em vista a sua importância em colocações com outros IL, muitas vezes contribuindo para que ele assuma a função de termo.

Segundo Maciel (2001):

“A nomenclatura do dicionário de língua geral – DLG – tende a abarcar a totalidade das palavras que compõem o léxico comum de uma língua, concentrando-se, de maneira especial, nas formas correntes na época de sua elaboração... Quanto a categorias

gramaticais, todas sem exceção estão representadas nas entradas. Unidades lexicais, selecionadas a partir de um *corpus* especializado, previamente estabelecido como representativo da área investigada, constituem, por outro lado, a nomenclatura do dicionário terminológico – DT – também conhecido como dicionário técnico ou dicionário técnico-científico... A categoria gramatical predominante é o substantivo”. [grifos meus]

#### **5.4.3. Candidatos a termo – decisão da autora**

Com base no acima exposto, a autora decidiu que o método a ser utilizado para a seleção da lista de entradas de um glossário técnico deve ser a listagem dos itens lexicais (doravante, ILs) portadores de sentido mais freqüentes, ou seja, os ILs que evocam uma idéia ou um conceito.

Na tabela a seguir estão listados os 100 IL mais freqüentes no *corpus* em língua portuguesa (doravante, LP). Os ILs mais freqüentes e portadores de sentido estão assinalados em negrito.

N.	WORD	FREQ.	%	N.	WORD	FREQ.	%
1	DE	5.807	5,30	51	CONSULTE	190	0,17
2	O	3.873	3,54	52	MANUTENÇÃO	190	0,17
3	A	3.605	3,29	53	RODAS	189	0,17
4	DO	2.853	2,60	54	PÁRA	186	0,17
5	E	2.238	2,04	55	TAMPA	185	0,17
6	PARA	1.638	1,50	56	SEMPRE	175	0,16
7	SE	1.285	1,17	57	TRASEIRO	175	0,16
8	EM	1.187	1,08	58	ÁGUA	174	0,16
9	COM	1.143	1,04	59	ALAVANCA	170	0,16
10	VEÍCULO	1.073	0,98	60	BANCO	166	0,15
11	DA	979	0,89	61	SEM	163	0,15
12	QUE	893	0,82	62	SUA	163	0,15
13	NO	798	0,73	63	LO	158	0,14
14	OS	795	0,73	64	PARTIDA	158	0,14
15	MOTOR	787	0,72	65	INTERRUPTOR	156	0,14
16	OU	713	0,65	66	QUANDO	156	0,14
17	NÃO	711	0,65	67	LUZ	155	0,14
18	AS	647	0,59	68	INDICADOR	154	0,14
19	DOS	573	0,52	69	LÂMPADA	153	0,14
20	É	543	0,50	70	BOTÃO	152	0,14
21	UM	542	0,49	71	CONTROLE	152	0,14
22	UMA	508	0,46	72	MARCHA	152	0,14
23	AR	482	0,44	73	BRISA	150	0,14
24	NA	481	0,44	74	DEVEM	149	0,14
25	À	479	0,44	75	APOS	148	0,14
26	SER	451	0,41	76	PARTE	146	0,13
27	SISTEMA	433	0,40	77	TEMPERATURA	145	0,13
28	AO	431	0,39	78	CINTO	143	0,13
29	DAS	396	0,36	79	CONCESSIONÁRIO	143	0,13
30	POSIÇÃO	396	0,36	80	LUZES	140	0,13
31	POR	348	0,32	81	PAINEL	140	0,13
32	SEGURANÇA	341	0,31	82	CINTOS	137	0,13
33	SEU	314	0,29	83	NUNCA	135	0,12
34	ÓLEO	282	0,26	84	ESTÁ	133	0,12
35	DEVE	278	0,25	85	UTILIZE	132	0,12
36	PNEUS	265	0,24	86	VIDROS	132	0,12
37	NÍVEL	260	0,24	87	MESMO	130	0,12
38	COMBUSTÍVEL	249	0,23	88	PRESSIONE	130	0,12
39	BATERIA	231	0,21	89	QUALQUER	130	0,12
40	IGNIÇÃO	230	0,21	90	ESTIVER	129	0,12
41	CASO	214	0,20	91	NECESSÁRIO	129	0,12
42	MAIS	209	0,19	92	PELO	128	0,12
43	FREIO	208	0,19	93	GIRE	127	0,12
44	PORTA	208	0,19	94	LADO	126	0,12
45	PORTAS	208	0,19	95	AIR	125	0,11
46	FUNCIONAMENTO	197	0,18	96	CARGA	125	0,11
47	DIREÇÃO	194	0,18	97	CONDIÇÕES	124	0,11
48	PODE	191	0,17	98	ANTES	123	0,11
49	ATÉ	190	0,17	99	VELOCIDADE	123	0,11
50	CHAVE	190	0,17	100	COMO	120	0,11

Tabela nº 2: 100 IL mais frequentes nos 5 documentos escritos em LP.

Em seguida, estão listados na tabela abaixo os 100 IL mais frequentes e portadores de sentido contidos no *corpus* em LP.

	N.	WORD	FREQ.	%		N.	WORD	FREQ.	%
1.	10	VEÍCULO	1.073	0,98	51.	106	MANUAL	116	0,11
2.	15	MOTOR	787	0,72	52.	107	RODA	115	0,10
3.	23	AR	482	0,44	53.	108	INSTRUMENTOS	113	0,10
4.	27	SISTEMA	433	0,40	54.	113	CAIXA	110	0,10
5.	30	POSIÇÃO	396	0,36	55.	114	BAG	109	0,10
6.	32	SEGURANÇA	341	0,31	56.	115	BANCOS	108	0,10
7.	34	ÓLEO	282	0,26	57.	117	MOVIMENTO	107	0,10
8.	36	PNEUS	265	0,24	58.	118	TRAÇÃO	106	0,10
9.	37	NÍVEL	260	0,24	59.	120	FARÓIS	104	0,09
10.	38	COMBUSTÍVEL	249	0,23	60.	121	FLUIDO	104	0,09
11.	39	BATERIA	231	0,21	61.	123	PRESSÃO	104	0,09
12.	40	IGNIÇÃO	230	0,21	62.	124	REBOQUE	104	0,09
13.	43	FREIO	208	0,19	63.	126	ARREFECIMENTO	101	0,09
14.	44	PORTA	208	0,19	64.	127	CONDUÇÃO	101	0,09
15.	45	PORTAS	208	0,19	65.	130	TRAVAMENTO	101	0,09
16.	46	FUNCIONAMENTO	197	0,18	66.	132	FORMA	99	0,09
17.	47	DIREÇÃO	194	0,18	67.	133	OBJETOS	99	0,09
18.	50	CHAVE	190	0,17	68.	134	COMANDOS	98	0,09
19.	51	CONSULTE	190	0,17	69.	136	UTILIZAÇÃO	97	0,09
20.	52	MANUTENÇÃO	190	0,17	70.	139	FAROL	96	0,09
21.	53	RODAS	189	0,17	71.	142	LIMPADOR	93	0,08
22.	54	PÁRA	186	0,17	72.	143	VOLANTE	93	0,08
23.	55	TAMPA	185	0,17	73.	151	FREIOS	90	0,08
24.	58	ÁGUA	174	0,16	74.	153	SUBSTITUIÇÃO	90	0,08
25.	59	ALAVANCA	170	0,16	75.	154	LÍQUIDO	85	0,08
26.	60	BANCO	166	0,15	76.	155	REGULAGEM	85	0,08
27.	64	PARTIDA	158	0,14	77.	156	TIPO	85	0,08
28.	65	INTERRUPTOR	156	0,14	78.	158	FILTRO	84	0,08
29.	67	LUZ	155	0,14	79.	159	CENTRAL	83	0,08
30.	68	INDICADOR	154	0,14	80.	161	ESPECIFICAÇÕES	80	0,07
31.	69	LÂMPADA	153	0,14	81.	163	PEDAL	80	0,07
32.	70	BOTÃO	152	0,14	82.	164	SENTIDO	80	0,07
33.	71	CONTROLE	152	0,14	83.	166	MACACO	79	0,07
34.	72	MARCHA	152	0,14	84.	169	EMBREGEM	77	0,07
35.	73	BRISA	150	0,14	85.	170	ALTURA	76	0,07
36.	76	PARTE	146	0,13	86.	173	PESO	76	0,07
37.	77	TEMPERATURA	145	0,13	87.	178	TROCA	75	0,07
38.	78	CINTO	143	0,13	88.	181	CONTATO	73	0,07
39.	79	CONCESSIONÁRIO	143	0,13	89.	182	DESLIGUE	73	0,07
40.	80	LUZES	140	0,13	90.	186	FLUXO	72	0,07
41.	81	PAINEL	140	0,13	91.	187	LÂMPADAS	72	0,07
42.	82	CINTOS	137	0,13	92.	189	TANQUE	72	0,07
43.	86	VIDROS	132	0,12	93.	190	VENTILADOR	72	0,07
44.	94	LADO	126	0,12	94.	191	ENCOSTO	71	0,06
45.	95	AIR	125	0,11	95.	192	ILUMINAÇÃO	71	0,06
46.	96	CARGA	125	0,11	96.	193	MODO	71	0,06
47.	97	CONDIÇÕES	124	0,11	97.	194	MOTORISTA	71	0,06
48.	99	VELOCIDADE	123	0,11	98.	195	VEÍCULOS	71	0,06
49.	103	VIDRO	118	0,11	99.	196	CRIANÇAS	70	0,06
50.	105	COMPARTIMENTO	116	0,11	100.	198	USO	70	0,06

Tabela nº 3: 100 IL portadores de sentido mais frequentes nos 5 documentos escritos em LP

Por meio da tabela acima, observamos que os quatro IL portadores de sentido mais frequentes no *corpus* em português são:

1. Veículo
2. Motor
3. Ar
4. Sistema



Dando continuidade às análises, foram listados os 100 IL mais frequentes nos 9 documentos constantes do *corpus* em LI. Em negrito estão destacados os ILs portadores de sentido.

N.	WORD	FREQ.	%	N.	WORD	FREQ.	%
1	THE	24.413	8,56	51	FUEL	700	0,25
2	TO	6.674	2,34	52	REAR	699	0,25
3	AND	6.606	2,32	53	FRONT	697	0,24
4	IS	4.207	1,47	54	SHOULD	690	0,24
5	IN	3.826	1,34	55	KEY	668	0,23
6	OR	3.724	1,31	56	ALL	653	0,23
7	A	3.555	1,25	57	IGNITION	652	0,23
8	OF	3.545	1,24	58	OIL	633	0,22
9	<b>VEHICLE</b>	<b>2.878</b>	<b>1,01</b>	59	SIDE	600	0,21
10	ON	2.850	1,00	60	PRESS	566	0,20
11	BE	2.346	0,82	61	<b>SAFETY</b>	558	0,20
12	FOR	2.218	0,78	62	LOCK	557	0,20
13	IF	2.159	0,76	63	UP	557	0,20
14	NOT	2.125	0,74	64	BAG	553	0,19
15	WHEN	1.913	0,67	65	SPEED	542	0,19
16	ENGINE	1.901	0,67	66	WHEEL	536	0,19
17	AIR	1.799	0,63	67	TIRE	532	0,19
18	WITH	1.777	0,62	68	ANY	530	0,19
19	SYSTEM	1.718	0,60	69	DEALER	528	0,19
20	IT	1.611	0,56	70	RESTRAINT	520	0,18
21	YOUR	1.598	0,56	71	ONLY	519	0,18
22	SEAT	1.365	0,48	72	FLUID	505	0,18
23	POSITION	1.314	0,46	73	MAINTENANCE	502	0,17
24	WILL	1.270	0,45	74	INFORMATION	499	0,17
25	THIS	1.265	0,44	75	WHILE	496	0,17
26	LIGHT	1.179	0,41	76	INDICATOR	495	0,17
27	MAY	1.158	0,41	77	AFTER	479	0,17
28	AS	1.142	0,40	78	BATTERY	470	0,16
29	ARE	1.138	0,40	79	LEVER	470	0,16
30	YOU	1.086	0,38	80	CD	466	0,16
31	DO	1.085	0,38	81	HAVE	465	0,16
32	SWITCH	1.066	0,37	82	CHECK	464	0,16
33	BUTTON	1.018	0,36	83	OPERATION	450	0,16
34	OFF	1.012	0,35	84	RADIO	449	0,16
35	USE	1.000	0,35	85	SECONDS	443	0,16
36	CONTROL	999	0,35	86	LIGHTS	442	0,15
37	BY	980	0,34	87	EQUIPPED	436	0,15
38	WARNING	948	0,33	88	CAUSE	433	0,15
39	THAT	939	0,33	89	LEVEL	426	0,15
40	FROM	906	0,32	90	DAMAGE	421	0,15
41	DRIVING	819	0,29	91	REMOVE	413	0,14
42	BELT	806	0,28	92	DISPLAY	412	0,14
43	S	792	0,28	93	SUPPLEMENTAL	411	0,14
44	AN	751	0,26	94	USED	409	0,14
45	TURN	751	0,26	95	BELTS	408	0,14
46	CAN	746	0,26	96	COULD	396	0,14
47	MODE	736	0,26	97	THAN	395	0,14
48	BRAKE	731	0,26	98	TEMPERATURE	394	0,14
49	AT	701	0,25	99	MUST	392	0,14
50	CHILD	700	0,25	100	DOOR	391	0,14

Tabela n° 4: 100 IL mais frequentes nos 9 documentos escritos em LI.

Em seguida, estão listados os 100 IL portadores de sentido mais frequentes constantes no *corpus* formado por 9 documentos redigidos em LI.

	N.	WORD	FREQ.	%		N.	WORD	FREQ.	%
1.	9	VEHICLE	2.878	1,01	51.	95	BELTS	408	0,14
2.	16	ENGINE	1.901	0,67	52.	98	TEMPERATURE	394	0,14
3.	17	AIR	1.799	0,63	53.	100	DOOR	391	0,14
4.	19	SYSTEM	1.718	0,60	54.	103	STEERING	377	0,13
5.	22	SEAT	1.365	0,48	55.	111	POWER	349	0,12
6.	23	POSITION	1.314	0,46	56.	112	COOLANT	346	0,12
7.	26	LIGHT	1.179	0,41	57.	114	PASSENGER	343	0,12
8.	32	SWITCH	1.066	0,37	58.	117	DRIVER	337	0,12
9.	33	BUTTON	1.018	0,36	59.	119	PEDAL	337	0,12
10.	34	OFF	1.012	0,35	60.	121	CONTROLS	334	0,12
11.	35	USE	1.000	0,35	61.	125	TRANSMISSION	326	0,11
12.	36	CONTROL	999	0,35	62.	127	DRIVE	325	0,11
13.	38	WARNING	948	0,33	63.	130	SECTION	314	0,11
14.	41	DRIVING	819	0,29	64.	133	SEATS	310	0,11
15.	42	BELT	806	0,28	65.	134	AUDIO	307	0,11
16.	45	TURN	751	0,26	66.	137	TYPE	300	0,11
17.	47	MODE	736	0,26	67.	139	START	288	0,10
18.	48	BRAKE	731	0,26	68.	143	TIME	281	0,10
19.	50	CHILD	700	0,25	69.	144	MOVE	279	0,10
20.	51	FUEL	700	0,25	70.	148	WINDOW	273	0,10
21.	52	REAR	699	0,25	71.	149	RELEASE	270	0,09
22.	53	FRONT	697	0,24	72.	150	CAUTION	268	0,09
23.	55	KEY	668	0,23	73.	152	MAKE	267	0,09
24.	57	IGNITION	652	0,23	74.	153	SHIFT	263	0,09
25.	58	OIL	633	0,22	75.	154	STATION	262	0,09
26.	59	SIDE	600	0,21	76.	155	WHEELS	262	0,09
27.	61	SAFETY	558	0,20	77.	156	CHANGE	259	0,09
28.	62	LOCK	557	0,20	78.	158	RESTRAINTS	254	0,09
29.	64	BAG	553	0,19	79.	160	STOP	249	0,09
30.	65	SPEED	542	0,19	80.	162	CONDITIONS	245	0,09
31.	66	WHEEL	536	0,19	81.	163	POSSIBLE	245	0,09
32.	67	TIRE	532	0,19	82.	164	DISC	244	0,09
33.	69	DEALER	528	0,19	83.	165	DOORS	244	0,09
34.	70	RESTRAINT	520	0,18	84.	166	TIRES	243	0,09
35.	72	FLUID	505	0,18	85.	168	PRESSURE	241	0,08
36.	73	MAINTENANCE	502	0,18	86.	169	PULL	241	0,08
37.	74	INFORMATION	499	0,17	87.	170	COMPARTMENT	240	0,08
38.	76	INDICATOR	495	0,17	88.	171	PANEL	239	0,08
39.	78	BATTERY	470	0,16	89.	172	OUTSIDE	234	0,08
40.	79	LEVER	470	0,16	90.	174	FUNCTION	233	0,08
41.	80	CD	466	0,16	91.	177	WATER	229	0,08
42.	82	CHECK	464	0,16	92.	178	NORMAL	227	0,08
43.	83	OPERATION	450	0,16	93.	179	FILTER	226	0,08
44.	84	RADIO	449	0,16	94.	181	NUMBER	225	0,08
45.	85	SECONDS	443	0,16	95.	183	HEATER	223	0,08
46.	86	LIGHTS	442	0,15	96.	184	PARKING	222	0,08
47.	88	CAUSE	433	0,15	97.	185	TYRE	219	0,08
48.	89	LEVEL	426	0,15	98.	186	INJURY	218	0,08
49.	90	DAMAGE	421	0,15	99.	189	INSTRUMENT	216	0,08
50.	92	DISPLAY	412	0,14	100.	192	BACK	211	0,07

Tabela nº 5: 100 IL portadores de sentido mais frequentes nos 9 documentos escritos em LI.

Através da tabela acima, observamos que os quatro IL portadores de sentido mais frequentes no *corpus* em inglês são:

1. Vehicle
2. Engine
3. Air
4. System

Nagao (1994) sugere que os termos de uma área do conhecimento são os nomes dados aos conceitos naquela área, ou seja, os substantivos. Por sua vez, Castillo (1998) concluiu que um termo é uma unidade denominativa de alguns dos membros do sistema conceitual de determinada área de especialidade. O termo não se distingue formalmente de qualquer outro signo e os fatores pragmáticos que o caracterizam são relativos. Assim, dentre os itens lexicais portadores de sentido listados tanto no *corpus* em LP quanto em LI, apenas os substantivos foram selecionados para se tornarem candidatos a termo no glossário técnico aqui em estudo. Por fim, o MT identifica que os conhecimentos adquiridos pelo terminólogo o ajudarão na hora de identificar a terminologia fundamental. Além de sua própria experiência profissional no ramo automobilístico, a autora contou, para a realização deste estudo, com a assessoria de um engenheiro mecânico, com especialização na área automobilística.

É importante ressaltar que este estudo trata da metodologia de criação de um glossário técnico e não da criação da obra, o que exigiria tanto o trabalho de profissionais da área de terminografia, quanto de revisão, edição e editoração, com dedicação exclusiva e integral dos profissionais envolvidos no projeto. No entanto, a autora sustenta que todos os termos obtidos através da lista de palavras originada com base no *corpus* em português devem ser relacionados no produto final, ou seja, o glossário de termos técnicos criado com base na metodologia proposta (inclusive as formas de frequência 1, ou *hapax legomena*). Através destas análises, percebe-se que, independentemente do país ou do idioma de origem de um manual destinado ao usuário de um veículo automotor, o seu vocabulário é basicamente constituído dos mesmos itens lexicais. Esta constatação confirma a proposição de se construir um glossário técnico-automobilístico através da utilização de um *corpus* formado por manuais

destinados aos proprietários de veículos automotores. Se a lista dos IL portadores de sentido se repete em diversos manuais, esta é então o ponto de partida para se criar um glossário, pois provavelmente estes seriam os ILs que os usuários do glossário mais provavelmente buscariam neste tipo de obra. Cabe ressaltar que a maior oferta de dicionários de especialidade contribui não só para a consulta por profissionais ligados a uma determinada área, mas também para auxiliar no trabalho dos tradutores técnicos. Ou seja, o objetivo principal de um dicionário de especialidade é a comunicação profissional. Atualmente, um dicionário de especialidade considerado satisfatório deve conter não apenas as equivalências entre dois ou mais idiomas, mas principalmente exemplos reais da utilização de cada item lexical, em cada idioma apresentado. Xatara (2001) adverte que, além do “fôlego de lexicógrafo”, é importante a visão de tradutor na árdua missão de elaborar um dicionário bilíngüe.

“Veículo” / “vehicle” e “motor” / “engine” são os dois substantivos mais freqüentes tanto no *corpus* em português quanto em inglês. Como o *corpus* aqui analisado se refere a manuais de utilização de veículos automotores, esta constatação parece razoável. Já “ar” / “air” é o terceiro substantivo mais freqüente no *corpus*. Castillo (1998) relata que alguns termos já nascem “terminologizados”. Outros se “terminologizam” em certo momento, quando se agrega outra acepção àquela conhecida dentro da língua geral, para denominar um conceito próprio de alguma área especializada, passando a ser usados como termos. Assim, um termo é uma unidade pragmática.

Em vista destas constatações, vamos examinar as ocorrências de “ar” e “air” no *corpus* de análise:

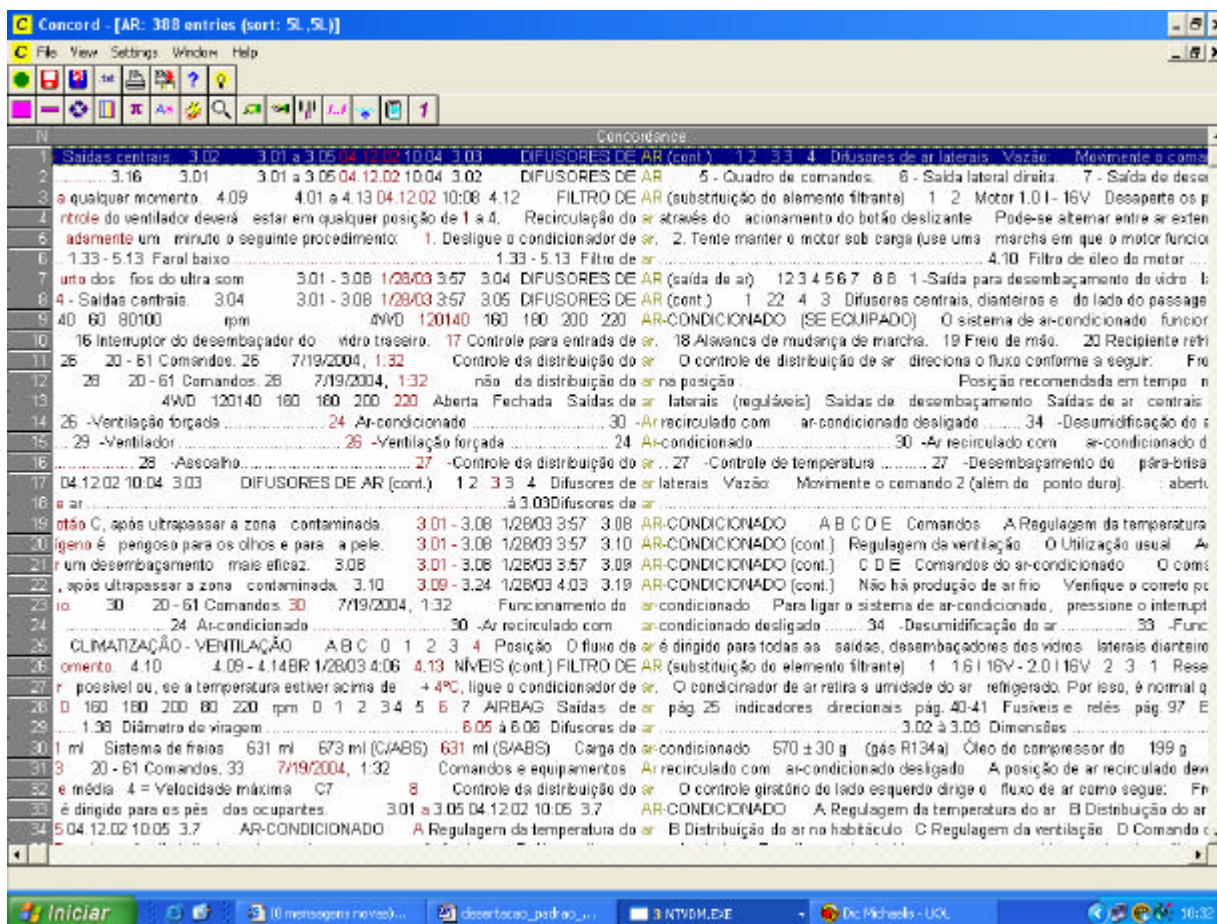


Figura nº 3: Ocorrências do IL “ar” no *corpus* em LP.

Analisando as ocorrências de “ar” no *corpus* em português, é possível extrair as seguintes colocações para o IL “ar”:

- Difusores de ar;
- Filtro de ar;
- Condicionador de ar;
- Saídas de ar (centrais ou laterais);
- Ar-condicionado;
- Circulador de ar;
- Difusor de ar;
- Ar insuflado;
- Ar recirculado;
- Ar frio;
- Ar quente;
- Condensador do ar condicionado;
- Interruptor de ar recirculado;
- Desumidificação do ar;

- Ar externo;
- Controle de distribuição de ar;
- Corrente de ar;
- Fluxo de ar;
- Pressão de ar;
- Indução de ar (aspirada ou forçada);
- Vazão do ar.

Para que se possa estabelecer uma comparação, são apresentadas as ocorrências do IL “air” no *corpus* em inglês.

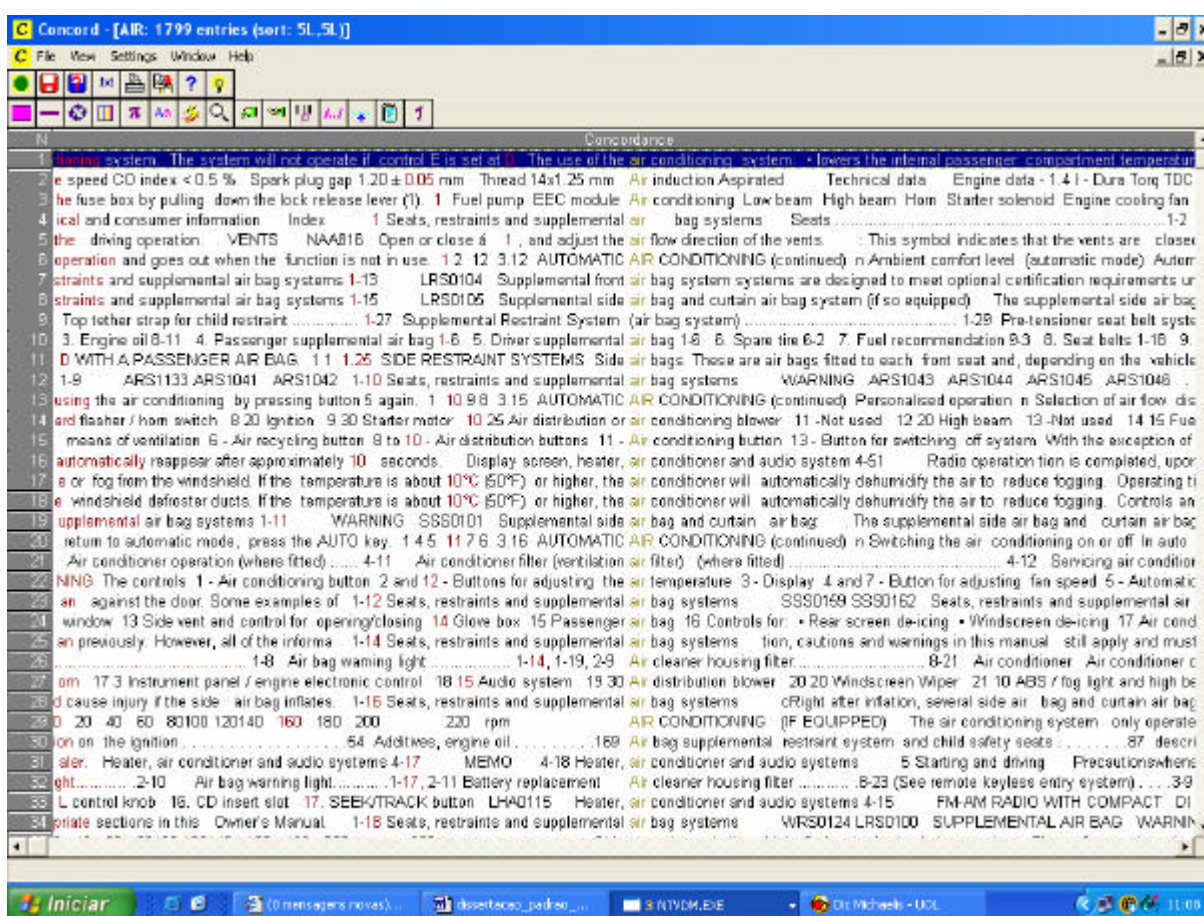


Figura nº 4: Ocorrências do IL “air” no *corpus* em LI.

Analisando as ocorrências de “air” no *corpus* em inglês, é possível extrair as seguintes colocações para o IL “air”:

- Air conditioning;
- Air conditioner;
- Air induction;

- Air bag;
- Air flow;
- Air filter;
- Air temperature;
- Air cleaner housing filter;
- Air distribution blower;
- Air outlets;
- Air vents;
- Sliding air recirculation control;
- Outside air;
- Fresh air mode;
- Outside air mode;
- Air bag location;
- Recirculated air;
- Air interchange;
- Air switch;
- Air cleaner;
- Air recirculation;
- Air recycling;
- Air inlet;
- Air adjusters;
- Air control;
- Air deflectors;
- Air filtration systems;
- Air intake area;
- Air release plug.

Além das numerosas ocorrências do termo “air bag”, que não tem correspondente em português, o IL “ar” aparece como substantivo principalmente em conjunto com “condicionado”. Nos demais casos, “ar” aparece acompanhado da preposição “de”, formando uma locução adjetiva, como em “saída de ar”, “entrada de ar”, “condicionador de ar”, “filtro de ar”, etc.

### **5.5. Extração de verbetes a partir do *corpus* de análise**

De acordo com o MT, para registrar o uso autêntico de uma especialidade, recomenda-se que primeiramente sejam recolhidos os termos dos textos originais



(língua de partida) e, depois, sejam recolhidos os termos de textos traduzidos (língua de chegada). A operação de extração de termos identifica não somente as unidades terminológicas, mas também os termos coocorrentes, por vezes chamados unidades fraseológicas, que ilustram o uso de um termo no discurso especializado.

De acordo com a lista de palavras mais frequentes no *corpus* de análise em questão, mencionada em 5.3 acima, tomaremos como exemplo o IL “alavanca”, relacionado com o número 25 na lista de palavras em português e com o número 79 na lista de palavras em inglês. Para isso, utilizaremos o manual do veículo “Ecosport”, que está disponível tanto em português quanto em inglês:

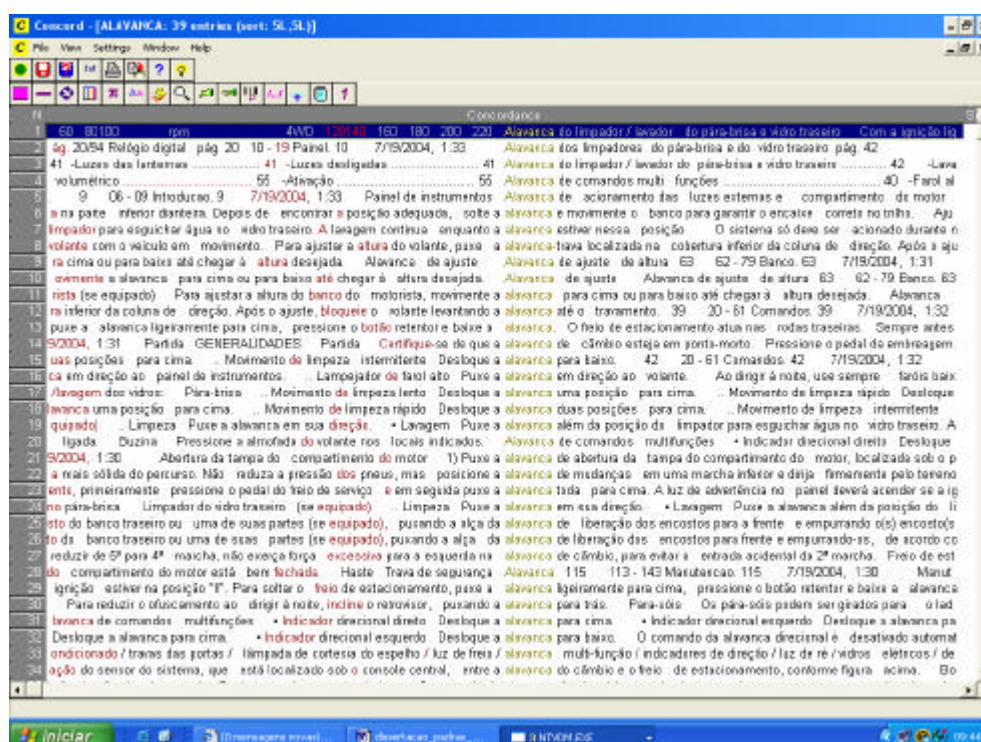


Figura nº 5: Ocorrências do IL "alavanca" no *corpus* em LP.

De acordo com a ocorrência nº 11, apresentada na lista de concordâncias do *corpus* em LP (figura nº 5), e a ocorrência nº 31, apresentada na lista de concordâncias



do *corpus* em LI (figura n° 6), podemos estabelecer a correspondência entre os ILs “alavanca” e “lever”.

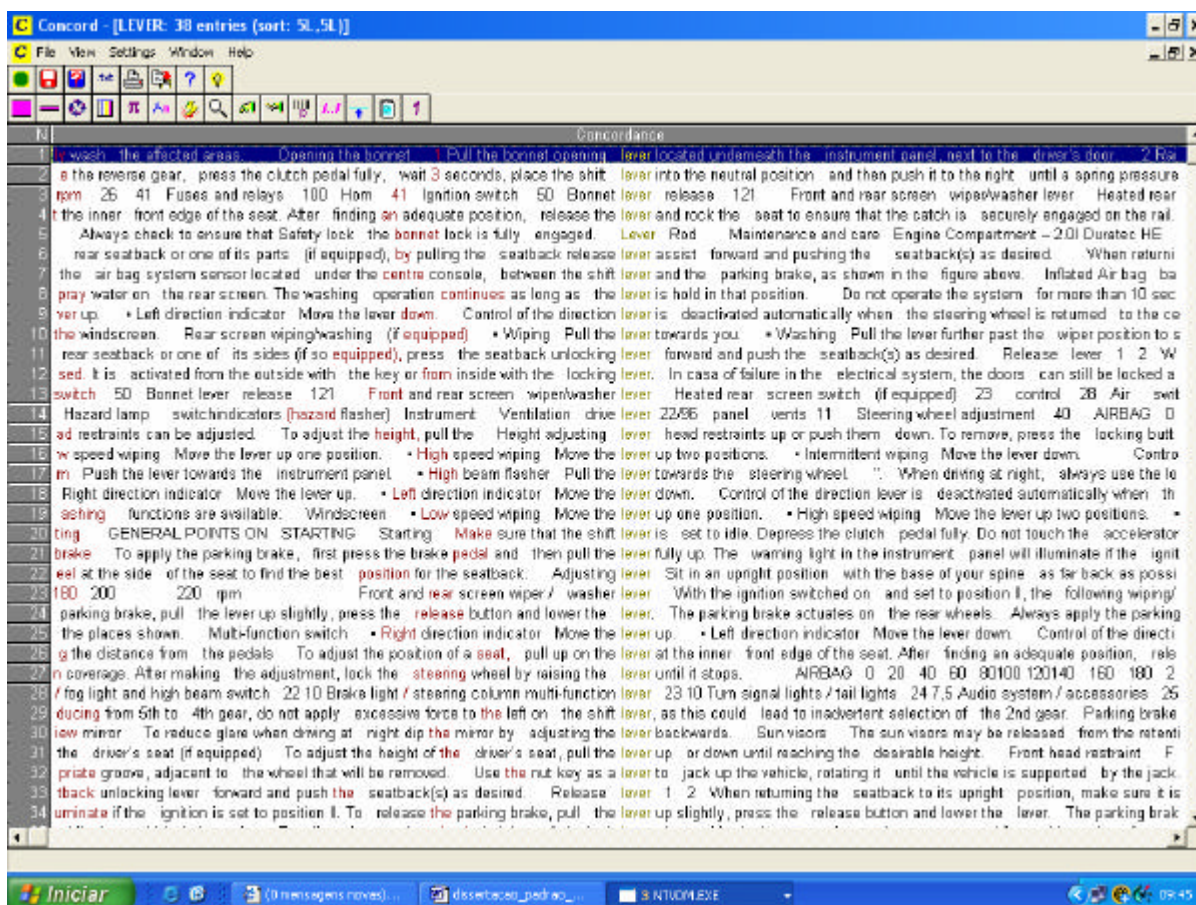


Figura n° 6: Ocorrências do IL "lever" no *corpus* em LI.

A partir das ocorrências acima apresentadas, podem ser extraídos os seguintes exemplos de verbetes para inclusão em um glossário técnico-automobilístico bilíngüe português / inglês – inglês / português:

Seção português → inglês	
<u>Português</u>	<u>Inglês</u>
<b>Alavanca</b> (s).	<b>Lever.</b> Ex.: To adjust the height of the driver's seat, pull the <i>lever</i> up or down until reaching the desirable height.

Tabela nº 6: Exemplo de verbete no glossário técnico-automobilístico bilíngüe português → inglês

Seção inglês → português	
<u>Inglês</u>	<u>Português</u>
<b>Lever</b> (n).	<b>Alavanca.</b> Ex.: Para ajustar a altura do banco do motorista, movimente a <i>alavanca</i> para cima ou para baixo até chegar à altura desejada.

Tabela nº 7: Exemplo de verbete no glossário técnico-automobilístico bilíngüe inglês → português

## 5.6. Formato

Ao considerar que vivemos em uma sociedade informatizada, recomenda-se que novas obras terminográficas sejam igualmente disponibilizadas em meio digital. No caso de um glossário técnico do ramo automobilístico, que poderá ser utilizado por

funcionários de empresas fabricantes de automóveis ou ainda de autopeças, deve-se pensar na publicação de um glossário tanto em meio digital, para agilizar a pesquisa pelos funcionários ditos “colarinhos brancos”, quanto uma versão impressa, para aqueles ditos “colarinhos azuis”, que normalmente não têm acesso a um computador em seu ambiente de trabalho. Independentemente do meio de divulgação e, parafraseando Rousseau (1998), o apressado cliente da Terminologia busca informações imediatas, precisas, concisas, que apresentem um alto grau de legibilidade e que estejam prontas para o uso. Este é o objetivo desta metodologia. Portanto, seja qual for o formato, a apresentação dos verbetes deve ser a mais clara e objetiva possível.

### 5.7. Sinonímia

“A existência do dicionário está baseada na hipótese, obviamente não comprovada, de que as línguas são compostas de sinônimos equivalentes”.

Jorge Luís Borges

A questão sobre a pertinência ou não do uso de sinônimos no que se refere à terminografia foi muito discutida. Entretanto, a melhor demonstração real não só de sua existência, como de sua utilidade, é o recurso Grand Dictionnaire Terminologique (francês ↔ inglês), do Office Québécois de la Langue Française, do Canadá. Buscando, por exemplo, o termo em francês *levier de frein*, que oferece como correspondente em LI *brake-actuating lever*, obtém-se, também, uma opção de sinônimo para o termo no idioma de partida (*levier de came de frein*) e quatro opções de quase-sinônimos (*levier de came*, *levier came*, *levier amplificateur d'effort*, *levier d'écartement des mâchoires de frein*). Entretanto, no idioma de chegada (ou seja, em LI), não é apresentado nenhum sinônimo para o termo e são apresentadas apenas três opções de quase-sinônimos (*brake*

control lever, cam lever, brake toggle lever). Castillo (1998) conclui que a sinonímia existe, em algumas áreas de especialidade com mais ou menos evidência do que em outras, e a nossa obrigação é registrá-la onde quer que esteja. Como a sua existência é inegável, o seu registro deve ser o mais exato possível, para não induzir o usuário a erros. Na metodologia aqui apresentada, preconiza-se a apresentação de sinônimos tanto para os termos do idioma de partida quanto de chegada, desde que identificados no *corpus* de análise.

No *corpus* analisado nesta pesquisa podem ser observados alguns casos de sinonímia, como será apresentado a seguir. Para isso, segue abaixo uma cópia da tela de visualização das ocorrências de “air conditioning” no *corpus* em LI.

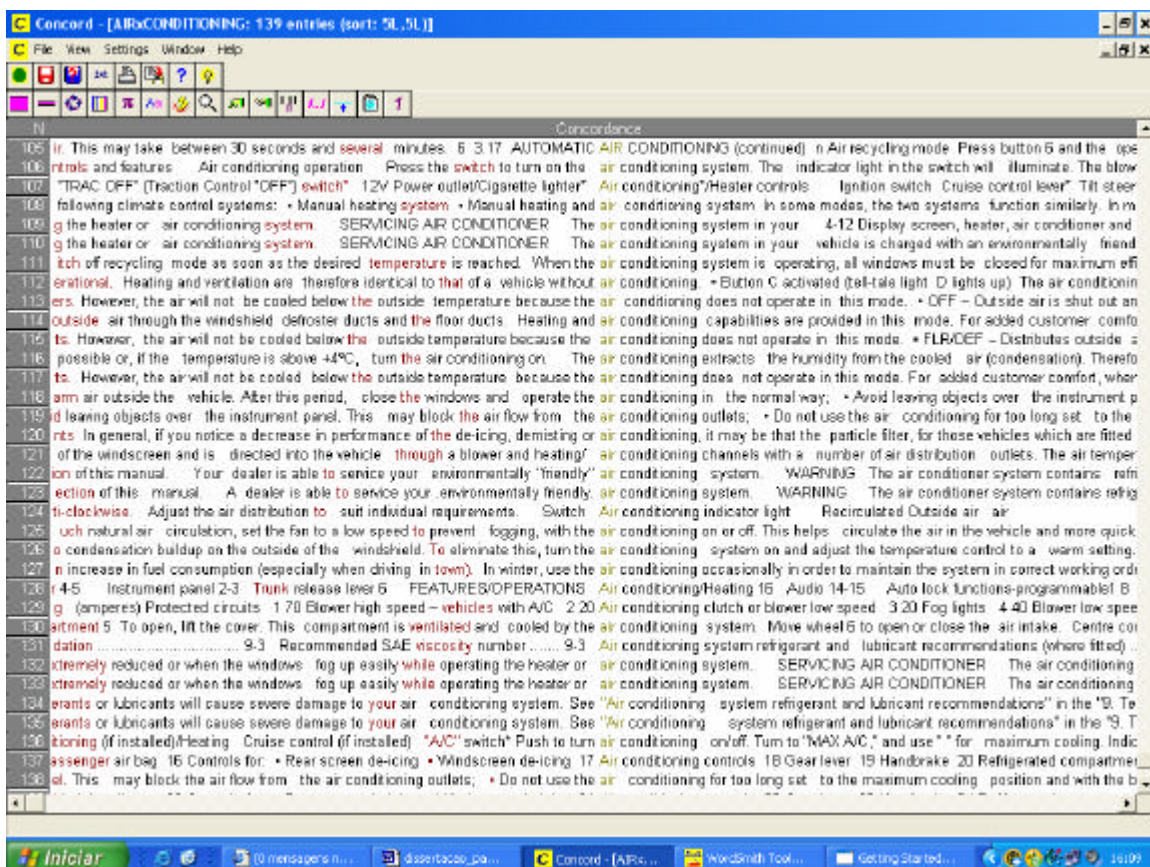


Figura nº 7: Ocorrências do IL “air conditioning” no *corpus* em LI.



No *corpus* em LI, foram listadas 139 ocorrências do termo “air conditioning”, em colocações como “Do not use the air conditioning for too long set to the maximum cooling position”. Vejamos, agora, quantas ocorrências são apresentadas para o termo “air conditioner”, no mesmo *corpus* em LI.

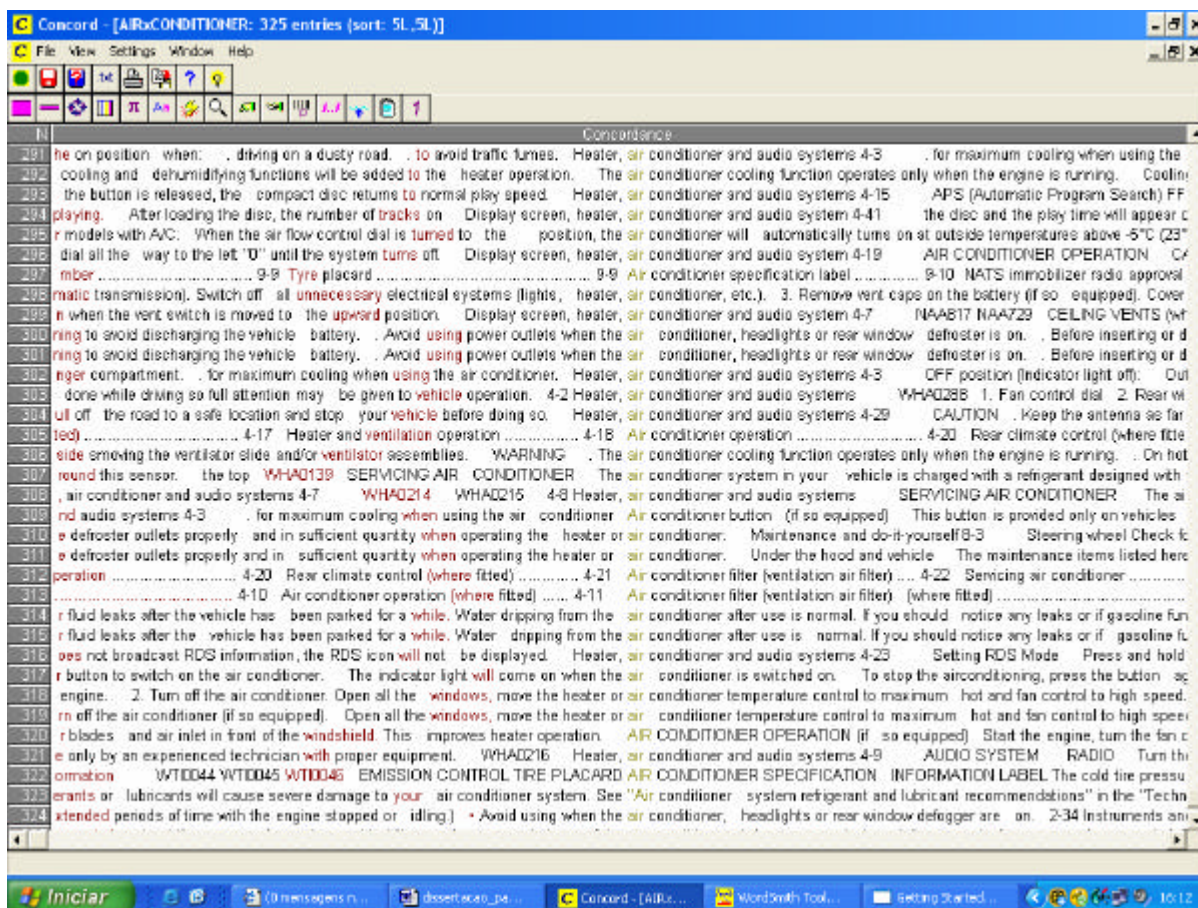


Figura nº 8: Ocorrências do IL “air conditioner” no *corpus* em LI.

No *corpus* em LI, encontramos um sinônimo para o termo “air conditioning”. “Air conditioner” aparece em 325 ocorrências, em colocações como “Avoid using when the air conditioner, headlights or rear window defogger are on”.

Vejamos, agora como ocorre o termo “ar-condicionado” no *corpus* em LP.

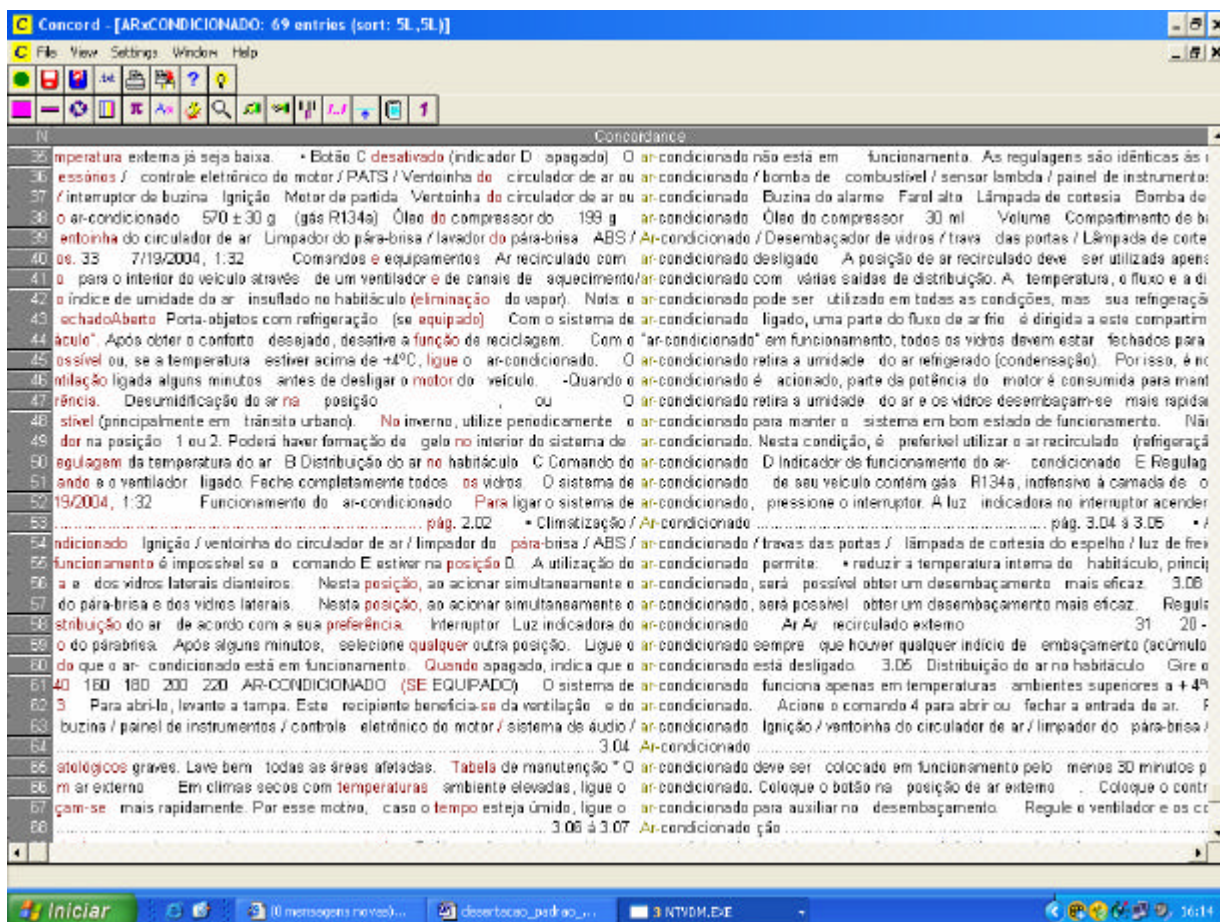


Figura nº 9: Ocorrências do IL “ar-condicionado” no *corpus* em LP.

“Ar-condicionado” aparece em 69 ocorrências, em colocações como “O ar-condicionado retira a umidade do ar refrigerado (condensação)”.

Segue agora uma visualização da tela de ocorrências do termo “condicionador de ar”, no *corpus* em LP.



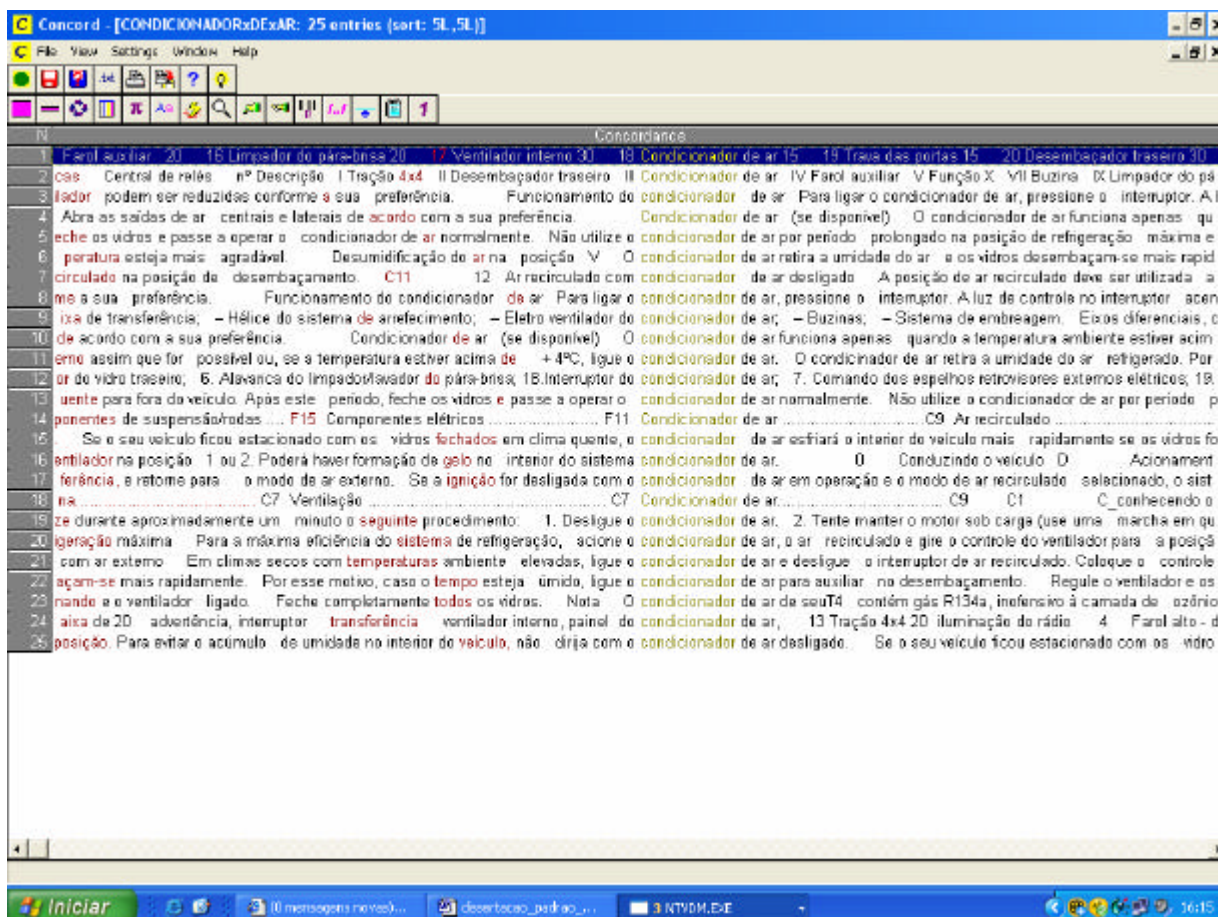


Figura nº 10: Ocorrências do IL “condicionador de ar” no *corpus* em LI.

Nas 25 ocorrências de “condicionador de ar”, sinônimo de “ar-condicionado”, podemos identificar colocações como “O condicionador de ar retira a umidade do ar”.

Tendo em vista as observações feitas em relação ao *corpus* em estudo, podemos verificar que existem duas equivalências tanto em LP quanto em LI para o termo que se refere ao “aparelho para regular a temperatura e a umidade de um ambiente fechado”<sup>10</sup>: “air conditioning” / “air conditioner”, “ar-condicionado” / “condicionador de ar”.

<sup>10</sup> Conforme Ferreira (1986).

### 5.7.1. Apresentação de sinônimos no glossário técnico

Como exemplo de apresentação de “air conditioner” / “air conditioning” / “ar-condicionado” / “condicionador de ar” no glossário técnico, podemos sugerir:

Seção português → inglês	
<u>Português</u>	<u>Inglês</u>
<b>Ar-condicionado</b> (s). <i>V. Condicionador de ar</i>	<b>Air conditioner / Air conditioning.</b>  Ex.: (1) The <i>air conditioner</i> cooling function operates only when the engine is running.  Ex.: (2) Close the windows and operate the <i>air conditioning</i> in the normal way.

Tabela nº 8: Exemplo de verbete no glossário técnico-automobilístico bilíngüe português → inglês

Seção inglês → português	
<u>Inglês</u>	<u>Português</u>
<b>Air conditioning</b> (n). <i>See Air conditioner</i>	<b>Ar-condicionado / Condicionador de ar.</b>  Ex.: (1) No inverno, utilize periodicamente o <i>ar-condicionado</i> para manter o sistema em bom estado de funcionamento.  Ex.: (2) Caso o tempo esteja úmido, ligue o <i>condicionador de ar</i> para auxiliar no desembaçamento.

Tabela nº 9: Exemplo de verbete no glossário técnico-automobilístico bilíngüe inglês → português



## 5.8. Ordem das Entradas

São possíveis dois tipos de ordem de classificação de entradas: pelo conteúdo ou pela forma (alfabética). Segundo Barros (2004), de acordo com o inconsciente coletivo do leitor, a ordem alfabética dos IL é a principal característica de um dicionário. Esta disposição simplifica a organização das entradas e a localização do IL pelo usuário. Para encontrar um verbete em um glossário organizado alfabeticamente, é necessário conhecê-lo antes. Como o objetivo do glossário bilíngüe não é apresentar a definição dos termos, mas sim a sua equivalência em dois idiomas e o registro dos respectivos exemplos reais de uso, esta forma de organização se confirma como sendo a melhor opção. Por isso, este foi o critério escolhido pela autora para a classificação das entradas.

## 5.9. Critérios de qualidade

De acordo com Rousseau (1998) e os princípios metodológicos da Realiter<sup>11</sup>, ao criar um banco de dados terminológico devem ser seguidos alguns critérios de qualidade, que são determinados sob o ponto de vista do usuário:

- **Acessibilidade:** o usuário deve encontrar rapidamente os dados terminológicos pertinentes às suas necessidades e adequados ao seu nível (nível de tecnicidade e de comunicação). A acessibilidade dos dados implica que eles devem ser apresentados em uma linguagem clara e compreensível. O usuário deve

---

<sup>11</sup> Princípios metodológicos do trabalho terminológico na Rede Panlatina de Terminologia – Realiter – 1995.

encontrar rapidamente a resposta às suas dúvidas no momento em que ele tiver necessidade.

- Atualidade: seja qual for o domínio ou área do conhecimento, o usuário deve ter acesso a uma terminologia atualizada.
- Confiabilidade: O usuário deve poder utilizar uma terminologia reconhecida pelos seus pares e pelo meio profissional ao qual ele se dirige.

### 5.10. Regionalismos

A diferenciação dos regionalismos deve ser demonstrada em um glossário terminológico levando-se em conta o público-alvo da obra. No caso desta metodologia, tem-se em mente um usuário brasileiro que deseje compreender termos técnicos que apresentem dissimilaridades em inglês americano ou britânico, que são as variantes possivelmente encontradas em documentos técnicos do segmento automobilístico.

Em vista disso, vamos mostrar a diferença entre o uso da palavra *electric*, no inglês britânico, em comparação à utilização da palavra *power*, no inglês americano. Esta análise é feita a partir da ferramenta *Concord* do WST.

Quando se faz referência a dispositivos acionados pelo condutor e ativados eletricamente pelo veículo, nota-se que, no inglês britânico, é utilizada a palavra *electric*. Os exemplos aqui demonstrados referem-se a **vidros elétricos** e **travamento elétrico das portas**.

Em primeiro lugar, serão apresentadas as ocorrências do IL “elétrico” (ou “elétrica” / “elétricos” / “elétricas”).

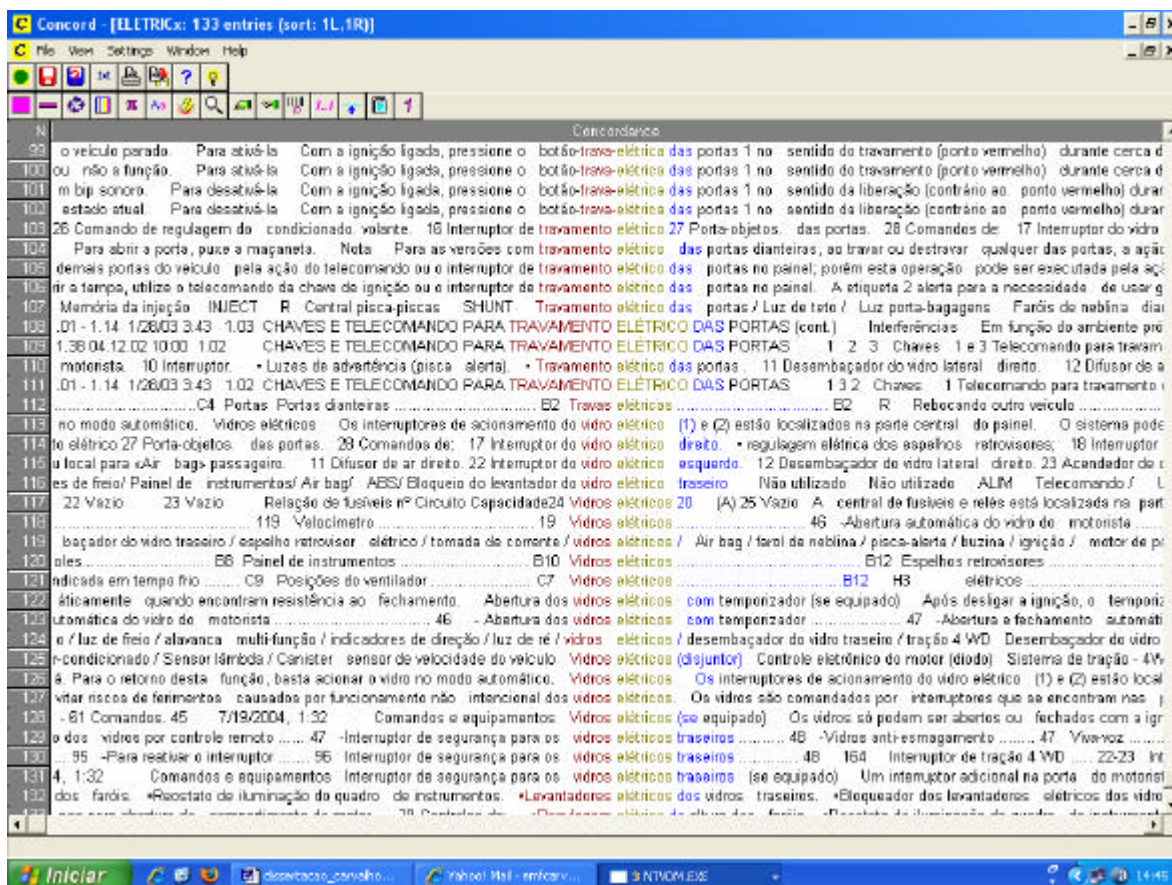


Figura nº 11: Ocorrências de “elétric\*” no *corpus* em LP.

Nas 133 ocorrências com “elétric\*” contidas no *corpus* em LP, encontramos exemplos como “Os interruptores de acionamento do vidro elétrico (1) e (2) estão localizados na parte central do painel”.

Agora, vejamos os exemplos com os ILs “electric” e “power” no *corpus* em LI. Em primeiro lugar, apresentaremos exemplos com o IL “electric”.

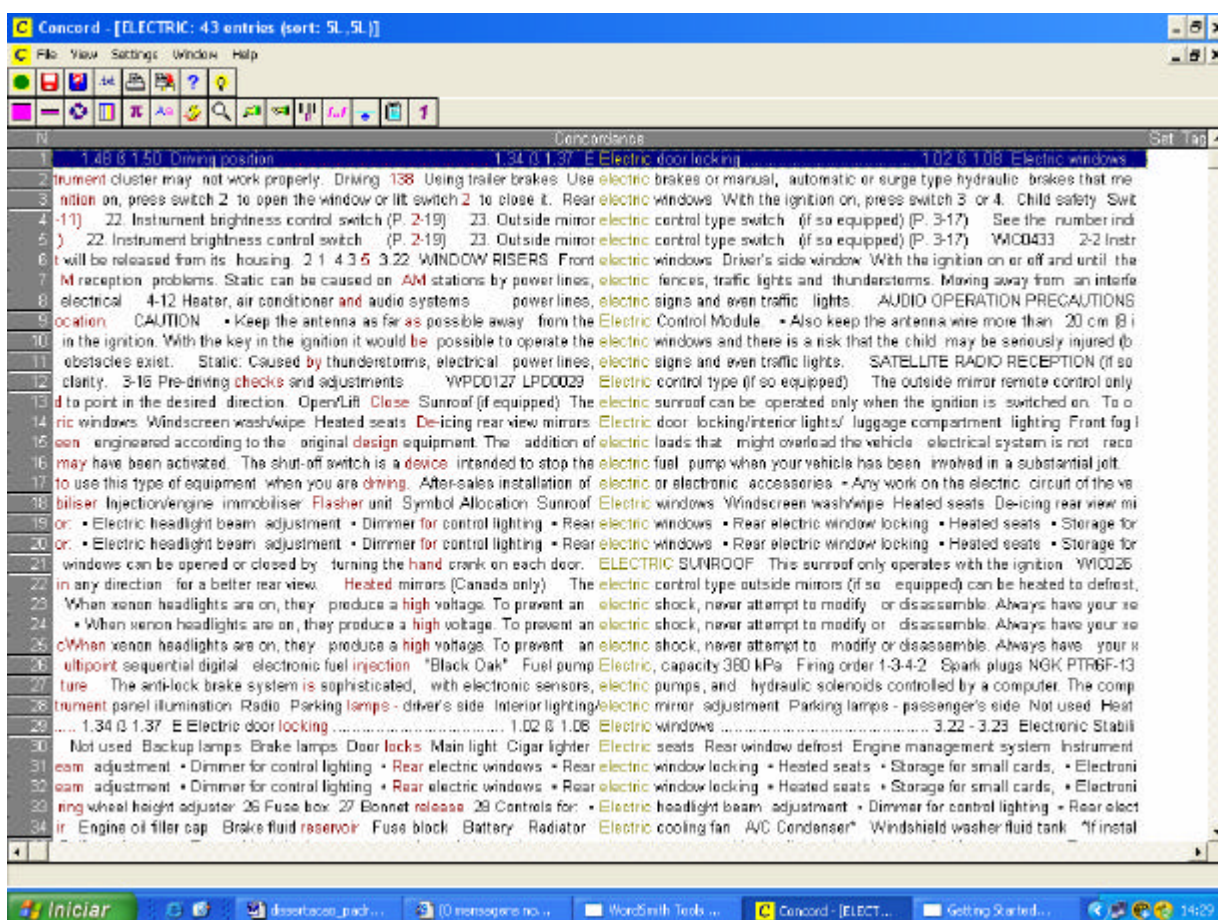


Figura nº 12: Ocorrências do IL “electric” no *corpus* em LI.

Nas 43 colocações de “electric”, podemos observar o uso deste IL em situações como “Whith the key in the ignition it would be possible to operate the electric windows and there is a risk that the child may be seriously injured”.

Quando a mesma referência é feita no manual escrito em inglês americano, a palavra utilizada é *power*:



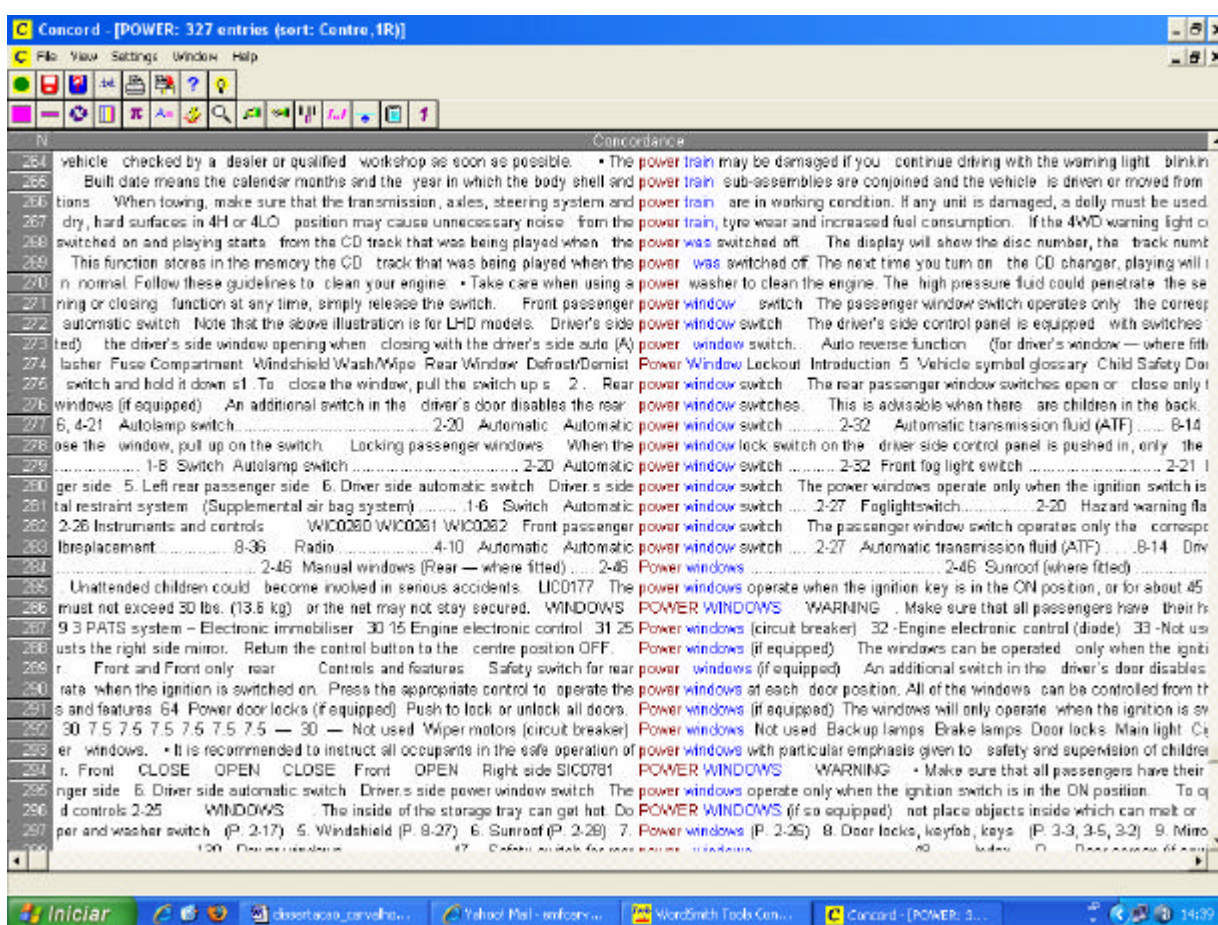


Figura nº 13: Ocorrências do IL “power” no *corpus* em LI.

No caso do emprego de “power” na variante americana da LI, observa-se, nas 349 ocorrências, exemplos como “An additional switch in the driver’s door disables the rear power window switches”.

Através das análises feitas no *corpus* em LI (total de 9 arquivos), e em LP (5 arquivos), temos os seguintes empregos dos IL “electric” e “power” significando “elétrico” / “elétrica” / “elétricos” / “elétricas” em português do Brasil:

Inglês (variante britânica)	Inglês (variante americana)	Português do Brasil
Rear electric window locker	Rear power window switches	Levantadores elétricos dos vidros traseiros
Electric door locking	Power door locks	Trava elétrica das portas

Tabela nº 10: Regionalismos

### 5.11. Possíveis diferenças

Até este ponto, todos os documentos foram analisados em conjunto, tendo sido agrupados apenas em relação ao idioma em que foram escritos, ou seja, LP e LI. Para dirimir eventuais dúvidas em relação ao conteúdo de textos comparáveis ou não, vamos analisar alguns documentos em separado, de forma a constatar eventuais diferenças em textos traduzidos ou originais, neste caso, utilizando o exemplo dos manuais escritos em LP.

Em relação ao aspecto abordado no item 5.4.3, "itens lexicais portadores de sentido", segue a lista dos cinco IL mais frequentes e portadores de sentido encontrados no Manual Troller:

- Veículo
- Motor
- Ar
- Posição
- Sistema

Comparando este documento escrito originalmente em português do Brasil com o *subcorpus* escrito em LP, há uma inversão de posição em relação aos IL "posição" e "sistema". Estes figuram respectivamente em 4º e 5º lugar no manual Troller. No *subcorpus* em LP, "sistema" aparece em quarto lugar e, conseqüentemente, "posição" aparece em quinto.

Em relação ao tema "sinonímia", abordado em 5.7, foi constatado que o termo "condicionador de ar", que figura 25 vezes no *corpus* em LP é, na verdade, o termo escolhido pela empresa Troller para representar o "aparelho elétrico que permite a manutenção de desejada temperatura num ambiente"<sup>12</sup>. Ao analisar o manual Troller,

---

<sup>12</sup> Conforme o Dic. Michaelis – UOL em CD, Melhoramentos, 1988

não foram encontradas ocorrências para o termo "ar-condicionado". Entretanto, foram encontradas exatamente 25 ocorrências para o termo "condicionador de ar".

## 6. RESULTADOS

Em relação ao objetivo da pesquisa, considera-se que o resultado foi atingido uma vez que, a partir de uma compilação de documentos de teor técnico, foi possível iniciar a criação de um banco de termos técnicos bilíngüe, com as suas respectivas correspondências nos idiomas português e inglês. Este é o primeiro passo para a criação de um glossário bilíngüe de termos técnicos, seja da área automobilística ou de qualquer outra área de especialidade.

O primeiro passo na elaboração do glossário foi possível ao se extrair verbetes com base nas listas de palavras relacionadas pelo programa WST. Enriquecendo a apresentação de cada entrada, foram fornecidos exemplos de utilização do termo apresentado no projeto de glossário, tanto para os verbetes em LP quanto para os verbetes em LI.

A apresentação de casos de sinonímia e regionalismos, respectivamente nos itens 5.7 e 5.10 desta dissertação, demonstram que não só é possível, como também relevante delinear as variações que um mesmo termo pode apresentar em uma determinada área do conhecimento, principalmente quando o *objeto* por ele representado faz parte do quotidiano de diferentes usuários de uma língua: engenheiros de projeto e de produção, operários da indústria automobilística, operários da indústria de autopeças, mecânicos de reparação automotiva e, finalmente, usuários de veículos. O mesmo se dá em relação ao emprego em diferentes países, como Estados Unidos e Inglaterra, ou Brasil e Portugal.

Um aspecto a ser salientado refere-se ao emprego de uma espécie de vocabulário-base no *corpus* analisado em LI e o *corpus* em LP. No item 5.4.3, foi visto



que os 4 IL portadores de sentido mais frequentes se repetem tanto no *corpus* em LP quanto no *corpus* em LI. Por outro lado, pode-se também constatar que há pequenas variações em relação a documentos redigidos originalmente em LP e textos traduzidos a partir da LI, conforme abordado em 5.11.

Em relação à pertinência da pesquisa na área de estudos da tradução, acredita-se que o objetivo também foi atingido, visto que tal método pode ser aplicado por outros pesquisadores na área da tradução utilizando uma coletânea de textos de qualquer área de especialidade. Como exemplo, poderíamos sugerir os seguintes temas:

- Engenharia elétrica: manual de equipamentos ou de motores elétricos;
- Engenharia eletrônica: manual de equipamentos eletrônicos, manual de assistência técnica de aparelhos celulares;
- Farmacêutica: bulas de medicamentos;
- Etc.

Por sua vez, tradutores que não ambicionam a pesquisa científica, mas que concentram a sua área de atuação em um determinado campo do saber, podem igualmente aplicar esta metodologia com vistas a organizar um banco de dados de termos técnicos, agilizando de tal modo o seu fluxo de trabalho. Neste sentido, considera-se pertinente o detalhamento referente ao preparo dos originais sugerido pela autora, conforme foi visto em 5.3, principalmente em relação ao agrupamento de vários arquivos de capítulos em um único arquivo.

O próprio glossário de termos técnicos, resultado desta metodologia, tem a sua aplicação assegurada no meio pertinente à engenharia automobilística: profissionais da indústria automobilística, desde engenheiros, técnicos e operadores, até os estudantes de

nível superior e técnico, a imprensa especializada, os professores de idiomas com foco no ensino especializado de idiomas.

## 7. CONSIDERAÇÕES FINAIS

Através deste estudo, foi possível demonstrar que um tradutor, um pesquisador ou até mesmo um estudante de Terminologia pode elaborar o seu próprio glossário de termos técnicos. Para isso, basta que ele reúna o seu *corpus* de um determinado campo do saber, faça os respectivos tratamentos nos originais, para que possam ser lidos em um concordanciador como o WST, e finalmente siga o procedimento de seleção de candidatos a termo.

A principal vantagem na utilização do concordanciador WST relaciona-se aos aspectos econômicos. A versão *freeware* pode ser facilmente obtida pela Internet. A sua utilização é bastante simples: basta ler o tópico de ajuda e executar as operações através dos ícones auto-explicativos exibidos pelo programa. Para alguns, um obstáculo possível seria em relação ao idioma: o programa WST está disponível apenas em inglês. Entretanto, Berber Sardinha (2004) apresenta o passo a passo para a utilização deste programa em português, facilitando enormemente o trabalho de pesquisadores e estudantes que não dominam a LI.

Hoje em dia existem vários programas de auxílio à tradução, conhecidos no mercado da tradução como *CAT – Computer-Aided Translation*, ou ferramentas de memória de tradução. Estes programas também possibilitam a criação de glossários personalizados, com base em *corpora* de tradução armazenados no computador do usuário. A maior limitação para a utilização destes programas, principalmente no Brasil, é o fator financeiro. O alto valor de mercado destes programas inviabiliza a sua utilização por um grande número de pesquisadores e estudantes de Terminologia e, em alguns casos, pelos próprios tradutores brasileiros. Isso confirma que a utilização do

WST se ajusta perfeitamente à realidade acadêmica brasileira, permitindo que qualquer pesquisador ou estudante de Terminologia possa criar o seu glossário de termos técnicos através da utilização de um programa disponível na Internet. O mesmo vale para tradutores técnicos que ainda não utilizam, ou não querem utilizar, pelo menos por enquanto, uma ferramenta de memória de tradução.

Esta pesquisa não tem como objetivo propor uma conclusão final, mas sim abrir caminho para outras linhas de pesquisa em Terminologia com base em *corpora* computadorizados. Pesquisadores, estudantes e tradutores podem ter informações detalhadas, em linguagem simples, sobre como organizar o seu próprio banco de dados de termos técnicos, sejam quais forem os idiomas de partida e de chegada e a área de estudo.

Na opinião da autora, o valor deste tipo de pesquisa se encontra na contribuição que ela oferece aos interessados nas aplicações práticas da Terminologia. Complementando comentários já abordados anteriormente nesta dissertação, há poucas publicações, principalmente em língua portuguesa, que apresentem pesquisas práticas relacionadas à Lingüística de *Corpus* com foco em Terminologia. A literatura disponível já aborda uma quantidade razoável de temas focando a Teoria da Terminologia. Assim, espera-se que outros pesquisadores produzam estudos de âmbito prático e que possam até mesmo contestar a metodologia aqui apresentada.

A associação da Lingüística de *Corpus* à Terminologia possibilitou a condução desta pesquisa. A autora espera que pesquisadores, estudantes e tradutores percebam a necessidade de estudo de ambas as disciplinas para conduzir uma pesquisa relacionada à Terminologia. Já foi citado nesta dissertação que, atualmente, estudos em Terminologia devem incluir a Lingüística de *Corpus*. Tendo a Lingüística de *Corpus* mudado a forma

como se estuda e se descreve uma língua, acredita-se que os estudos em Terminologia devam seguir a mesma direção. Assim, a terminografia, como disciplina relacionada ao registro dos termos, deve se basear em dados coletados através do seu uso real pelos falantes nativos de um idioma, principalmente no que se refere aos exemplos.

## 8. REFERÊNCIAS

ALBUQUERQUE, Y. J. D. M. & CERCEAU R. "Implantação de um Controle Terminológico numa Indústria Siderúrgica: O Caso CSN". Actas del V Simposio Iberoamericano de Terminología RITerm. Cidade do México, 1996. Disponível em <http://www.riterm.net/actes/5simposio/maranha2.htm>

ALLEN, Robert E. A Concise History of the COD (Concise Oxford Dictionary). In: R. R. K. Hartmann (ed.) *Amsterdam studies in the theory and history of linguistic science. Series III, Studies in the history of language sciences*, Amsterdã / Filadélfia, John Benjamins, 1986.

ALPÍZAR CASTILLO, Rodolfo. Ideas sobre el trabajo terminográfico. In: Maria Helena Mateus e Margarita Correa (org.) *Terminologia: questões teóricas, métodos e projectos*. Mem Martins, Europa-América, 1998.

BAKER, Mona. *Corpora in translation studies: an overview and some suggestions for future research*. John Benjamins, 1995 (Target 7.2, 223-243).

BARROS, Lidia Almeida. *Curso Básico de Terminologia*. São Paulo. Editora da Universidade de São Paulo, 2004.

BERBER SARDINHA, Tony. Lingüística de *corpus*: história e problemática. DELTA. 2000, vol. 16, no. 2. Disponível em: [http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S0102-44502000000200005&lng=en&nrm=iso](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0102-44502000000200005&lng=en&nrm=iso).

BERBER SARDINHA, Tony. *Lingüística de Corpus*. Barueri, Manole, 2004.

FERREIRA, Aurélio Buarque de Holanda. *Novo dicionário da língua portuguesa*. Rio de Janeiro, Nova Fronteira, 1986.

HARVEY, K. e D. YUILL. *The COBUILD testing initiative: The introspective, encoding component*. Artigo científico não publicado. Cobuild, University of Birmingham, 1994.

KRIEGER, Maria da Graça e FINATTO, Maria José Bocorny. *Introdução à terminologia: teoria e prática*. São Paulo, Contexto, 2004.

MACIEL, A. M. B. Da análise e reconhecimento de candidatos a termo de um dicionário jurídico. In: Maria da Graça Krieger e Anna Maria Becker Maciel (orgs.) *Temas de terminologia*. Porto Alegre / São Paulo, Universidade / UFRGS / Humanitas / USP, 2001.

NAGAO, Makoto. A methodology for the construction of a terminology dictionary. In: B.T.S. Atkins e A. Zampolli (eds.) *Computational approaches to the lexicon*. Nova York, Oxford, 1994.

PARTINGTON, Alan. *Patterns and meanings: using corpora for English language research and teaching*. Amsterdã, John Benjamins, 1998.

PAVEL, S. e NOLET, D. *Manual de Terminologia (on-line)*. Tradução de Enilde Faulstich. Disponível em [http://www.translationbureau.gc.ca/pwgsc\\_internet/en/publications/documents/presport.pdf](http://www.translationbureau.gc.ca/pwgsc_internet/en/publications/documents/presport.pdf). Arquivo capturado em 11 de agosto de 2005.

ROUSSEAU, Louis-Jean. Les pratiques terminographiques en contexte de développement d'une banque de terminologie. In: Maria Helena Mateus e Margarita Correa (org.) *Terminologia: questões teóricas, métodos e projectos*. Mem Martins, Europa-América, 1998.

SCOTT, Mike. *WordSmith Tools*. Version 3.0, 1999, disponível em [www.liv.ac.uk/~ms2928](http://www.liv.ac.uk/~ms2928).

SVENSÉN, Bo. *Practical lexicography: principles and methods of dictionary-making*. Tradução do sueco para o inglês por John Sykes e Kerstin Schofield. Oxford / Nova Iorque, Oxford, 1993.

TAGNIN, Stella. *Os corpora. Instrumentos de auto-ajuda para o tradutor*. Cadernos de Tradução n° 9, Universidade Federal de Santa Catarina, disponível em <http://www.cadernos.ufsc.br/>, acessado em 24/11/06.

XATARA, Cláudia Maria. Os dicionários bilíngües e o problema da tradução. In: Ana Maria Pinto Pires de Oliveira e Aparecida Negri Isquierdo (org.) *As ciências do léxico: lexicologia, lexicografia, terminologia*. Campo Grande, UFMS, 2001.



“Toutes les autres auteurs peuvent aspirer à la louange; les lexicographes ne peuvent  
aspirer qu’à échapper aux reproches”.

(Ferreira, apud Andrieux, apud Alfred Elwall, Dictionnaire français -anglais)