

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

**O USO DE CORPORA COMO FERRAMENTA  
DE APOIO PARA TRADUÇÃO: Uma análise das  
co-ocorrências do item lexical “hearing”**

**JANE MARIAN**

**Florianópolis, julho de 2010**

JANE MARIAN

**O USO DE CORPORA COMO FERRAMENTA  
DE APOIO PARA TRADUÇÃO: Uma análise das co-ocorrências  
do item lexical “hearing”**

*Dissertação apresentada ao curso de  
Mestrado do Programa de Pós-  
Graduação em Estudos da Tradução  
da Universidade Federal de Santa Ca-  
tarina como requisito para a obtenção  
do título de Mestre em Estudos da Tra-  
dução sob a orientação do Prof. Dr.  
Marco Rocha.*

**Florianópolis, julho de 2010**

Dissertação julgada para a obtenção de grau de  
MESTRE EM ESTUDOS DA TRADUÇÃO  
Área de Concentração: Processos e Retextualização

**Lexicografia, Tradução e Ensino de Línguas**

Aprovada em sua forma final pelo programa de  
Pós Graduação em Estudos da Tradução da  
Universidade Federal de Santa Catarina – UFSC

Data de aprovação: \_\_\_\_ de \_\_\_\_\_ de 20\_\_\_\_.

---

Prof. Dra. Maria Lúcia Vasconcellos (UFSC)  
(Banca examinadora)

---

Prof. Dra. Tânia Maria Granja Shepherd (UERJ)  
(Banca examinadora)

---

Prof. Dr. Lincoln Fernandes (UFSC)  
(Banca examinadora)

---

Prof. Dr. Marco Rocha (UFSC)  
(Orientador)

---

Prof. Dr. Walter Carlos Costa (UFSC)  
(Coordenador)

*“The largest of the accomplishments is to live life doing what matters. Learning and teaching what you have learned are important in my view. So, I dedicate this work to all people who in one way or another contributed to my learning process” Jane Marian.*

## AGRADECIMENTOS

Ao meu orientador, Prof. Dr. Marco Rocha pelo exemplo, paciência e amizade.

Aos demais professores da PGET e PGI, que sempre demonstraram amizade e interesse em ajudar nas dúvidas existentes até o presente em especial aos professores Lincoln Fernandes, Maria Lúcia Vasconcellos, Ronaldo Lima e Markus J. Weininger.

À minha mãe e ao meu pai pelo apoio e suporte emocional, carinho e incentivo que sempre me deram para estudar.

E aos meus amigos sinceros que desde o início estiveram presentes na minha vida contribuindo de uma forma ou outra para que eu pudesse realizar este projeto.

E especialmente a Ilan Thomé da Silva pelos incentivos diários que precisei durante o percurso e pelo enorme carinho e amizade que tem demonstrado durante todos os dias desde novembro de 1998, quando começamos a namorar.

*“Keep away from people who try to belittle your ambitions. Small people always do that, but the really great make you feel that you, too, can become great”. Mark Twain.*

## RESUMO

O objetivo desta dissertação é o estudo do uso de *corpora* como ferramenta de apoio para a tradução. A proposta é fazer uma análise das palavras que co-ocorrem ao redor do item lexical *hearing* e quais as possíveis traduções em dois *corpora* diferentes: um dos *corpora* que será utilizado na pesquisa é o *Compara*, que possui textos literários, e o outro é o *EuroParl*, com textos transcritos da área legislativa. O interesse pelo estudo de *corpora* como ferramenta de apoio para traduções se faz devido às dificuldades que o tradutor tem em encontrar materiais auxiliares para a tradução, principalmente quando se trata de uma área especializada. Infelizmente os *corpora* paralelos especializados disponíveis gratuitamente na *web* também são poucos. No presente trabalho serão analisados os resultados encontrados na busca da palavra *hearing* em um *corpus* especializado da área jurídica/legislativa *versus* um *corpus* especializado da área literária.

**Palavras-chave:** Estudos da Tradução. Linguística de Corpus. Corpus especializado. Hearing. Co-ocorrência.

*ABSTRACT*

The objective of this dissertation project is the study of the *corpora* use as a support tool for translation. The proposal is to analyze the words that co-occur around the lexical item *hearing* and their possible translations in two different *corpora*: one of these *corpora* that will be used in the research is *Compara* that possesses literary texts and the other one is *EuroParl* with transcribed texts from legislative area. The interest for the *corpora* study as a support tool for translations is made due to the difficulties that the translator has in finding materials auxiliaries mainly for technical translation terms. Unfortunately free specialized parallel *corpora* available in the *web* are also few. The present study will be explored which would be the results found in the search of the word *hearing* in a specialized *corpus* of the legal/legislative area *versus* a specialized *corpus* of the literary area.

**Key words: Translation Studies. Corpus Linguistics. Juridical and Literary Translation. Especialized Corpus Hearing. Co-occurrence.**

## ***LISTA DE FIGURAS***

Figura 1 – Holmes`s map of translation (from Toury 1995:10).....	21
Figura 2 – Mapeamento sugerido pela autora (Marian, 2010).....	23
Figura 3 – Interface Natools.....	34
Figura 4 – Interface Compara.....	39
Figura 5 – Interface WordSmith Tools.....	44

## ***LISTA DE TABELAS***

Tabela 1 – Tamanho do Corpus .....	28
Tabela 2 - Lista de Co-ocorrências do Item Lexical Hearing (Corpus EuroParl).....	46
Tabela 3 – Correspondente Tradutório (total) do Item Lexical Hearing (Corpus EuroParl) .....	62
Tabela 4 – Co-ocorrêntes do Item Lexical hearing (Corpus Compara.....	63
Tabela 5 – Correspondentes Tradutórios do Corpora.....	64

## ***LISTA DE QUADROS***

Quadro 1 – Tipologia do Corpus (síntese da autora) .....	26
Quadro 2 – Corpora Paralelos (uma amostra dos Corpora Compara e EuroParl).....	40
Quadro 3 – Ocorrências e co-ocorrências do item lexical during + hearing (Corpus EuroParl).....	48
Quadro 4 – . Correspondentes Tradutórios do item lexical during + hearing (Corpus EuroParl).....	49
Quadro 5 – Ocorrências e co-ocorrências do item lexical after + hearing (Corpus EuroParl).....	50
Quadro 6 – Correspondentes Tradutórios do item lexical after + hearing (Corpus EuroParl).....	51
Quadro 7 – Ocorrências e co-ocorrências do item lexical held + hearing (Corpus EuroParl).....	52
Quadro 8 – Correspondentes Tradutórios do item lexical held + hearing (Corpus EuroParl) .....	53

Quadro 9 – Ocorrências e co-ocorrências do item lexical public + hearing (Corpus EuroParl).....	54
Quadro 10 – Correspondentes Tradutórios do item lexical public + hearing (Corpus EuroParl). ....	55
Quadro 11 – Ocorrências e co-ocorrências do item lexical hearing + about (Corpus EuroParl). ....	56
Quadro 12 – Correspondentes Tradutórios do item lexical hearing + about (Corpus EuroParl). ....	56
Quadro 13 – Ocorrências e co-ocorrências do item lexical look forward to hearing (Corpus EuroParl) .....	57
Quadro 14 – Correspondentes Tradutórios do item lexical look forward to hearing (Corpus EuroParl) .....	58
Quadro 15 – Ocorrências e co-ocorrências do item lexical on hearing e hearing on (Corpus EuroParl). ....	59
Quadro 16 – Correspondentes Tradutórios do item lexical on hearing e hearing on (Corpus EuroParl).....	59
Quadro 17 – Ocorrências e co-ocorrências do item lexical from hearing e hearing from (Corpus EuroParl).....	60
Quadro 18 – Correspondentes Tradutórios do item lexical from hearing e hearing from (Corpus EuroParl).....	61
Quadro 19 – Co-ocorrências do item lexical hearing + verb to be (Corpus Compara).....	65
Quadro 20 – Correspondentes Tradutórios do item lexical hearing + verb to be (Corpus Compara) .....	65
Quadro 21 – Co-ocorrências do item lexical on hearing e hearing on (Corpus Compara). ....	66
Quadro 22 – Correspondentes Tradutórios do item lexical on hearing e hearing on (Corpus Compara).....	67
Quadro 23 – Co-ocorrências do item lexical hard of hearing (Corpus Compara).....	68
Quadro 24 – Correspondentes Tradutórios do item lexical hard of hearing (Corpus Compara).....	69
Quadro 25 – Co-ocorrências do item lexical in hearing (Corpus Compara).....	69
Quadro 26 – Correspondentes Tradutórios do item lexical in hearing (Corpus Compara). ....	70
Quadro 27 – Co-ocorrentes do Corpus EuroParl versus Compara.....	71
Quadro 28 – Correspondentes Tradutórios do Corpus EuroParl versus Compara.....	73

## SUMÁRIO

INTRODUÇÃO.....	13
Organização do Trabalho.....	17

### CAPÍTULO 1

1. REVISÃO DA LITERATURA.....	19
1.1. Definição de Corpus.....	19
1.2. Qual a finalidade de um corpus?.....	20
1.3. Estudos da tradução.....	21
1.4. Corpora em Estudos da Tradução.....	24
1.4.1. Corpora Paralelos.....	25
1.4.2. Corpora Comparáveis.....	25
1.4.3. Corpora Multilíngues.....	25
1.4.4. Classificação do Corpus.....	26
1.5. Linguística de Corpus.....	28
1.6. A Importância da Terminologia na Tradução.....	32

### CAPÍTULO 2

2. METODOLOGIA.....	33
2.1. Desenho do Corpus EuroParl.....	33
2.2. Desenho do Corpus Compara.....	36
2.3. Procedimentos de coleta da amostra do item lexical <i>hearing</i> .....	39
2.4. Procedimentos para Processamento do Corpus.....	43
2.5. Procedimentos para Análise de Dados.....	44

## CAPÍTULO 3

3.	ANÁLISE DOS DADOS.....	45
3.1.	Análise das co-ocorrências do item lexical <i>hearing</i> e suas respectivas traduções no Corpus EuroParl.....	45
3.2.	Análise das co-ocorrências do item lexical <i>hearing</i> e suas respectivas traduções no Corpus Compara.....	46
3.3.	Análise dos Resultados Gerais.....	62
4.	CONSIDERAÇÃO FINAIS.....	76
5.	REFERÊNCIAS.....	78
6.	ANEXOS.....	82

## INTRODUÇÃO

Esta dissertação descreve uma pesquisa que surgiu a partir de interesses pessoais em desenvolver habilidades de tradução por meio do suporte tecnológico, em especial o uso de *corpus* e a curiosidade de entender como se procedem os estudos para o desenvolvimento de dicionários, tanto os terminológicos quanto os lexicográficos.

A presente pesquisa se dedica ao estudo do item lexical “*hearing*” em dois *corpora*: (i) o *EuroParl*, um *corpus* legislativo com 28.521.967 palavras em inglês e 29.213.348 palavras em português, sendo que possui textos originais em inglês com suas traduções para o português, textos originais em português com suas traduções para o inglês e textos em uma terceira língua usada no Parlamento Europeu (podendo ser qualquer uma das línguas oficiais usadas nesta organização) que são traduzidas para o inglês e português; e o (ii) *Compara*, um *corpus* literário, também alinhado, paralelo, bidirecional de português e inglês, com textos originais nas duas línguas e suas respectivas traduções, com 1.435.926 palavras em português (originais e traduções), e 1.542.762 palavras em inglês (originais e traduções).

A ideia de comparar dois *corpora* de áreas diferentes surgiu devido a estes motivos específicos: (i) durante um curso preparatório de tradução juramentada, ministrado pelas tradutoras juramentadas Lúcia Nascimento e Rosemary I. Carvalho Teixeira, no qual elas desenvolveram um glossário em que o item lexical *hearing* apresentava apenas a acepção *audiência* como correspondente tradutório da área jurídica; (ii) nas aulas do professor Marco Rocha aprendi a usar *corpora* para análises linguísticas e fui apresentada ao *corpus Compara* e a ferramenta computacional *WordSmith Tools*, que auxilia na análise das co-ocorrências do item lexical de estudo; e (iii) nas aulas do professor Lincoln Fernandes conheci a ferramenta *NaTools* e os *corpora EuroParl* que me levaram a investigar o item lexical *hearing* e comparar com os resultados do *Compara*.

A partir das análises e investigações em sala de aula iniciaram-se as leituras sobre o tema e, mais tarde, a decisão efetiva de apresentar uma metodologia baseada no uso de *corpora* como ferramenta de apoio para a tradução. Para demonstrar esta metodologia optei por explorar o item lexical *hearing* e analisar como ele se comporta dentro dos dois *corpora* com contextos diferentes, observando suas co-ocorrências e possíveis traduções.

O interesse pelo estudo de *corpora*, como ferramenta de apoio às traduções, se faz devido à falta de dicionários com exemplos reais, principalmente, aos tradutores de diversas áreas técnicas. Um dos dicionários mais conhecidos na área jurídica é o Dicionário Jurídico Noronha (Noronha's Legal Dictionary), que apresenta como correspondente tradutório para o item lexical *hearing* apenas a acepção *audiência*, assim como no glossário elaborado pelas duas tradutoras juramentadas.

No entanto, ao observar o *corpus EuroParl*, percebemos que *hearing* apresenta mais de um correspondente tradutório dependendo do contexto apresentado. Com isto, algumas perguntas de pesquisa surgiram durante o processo de investigação: (i) Como se comporta o item lexical “*hearing*” nos contextos legislativo e literário?; (ii) As palavras que co-ocorrem com “*hearing*” são diferentes nos dois contextos?; (iii) Qual a importância do *corpus* para averiguar semelhanças e diferenças de um item lexical?; (iv) Que influência as co-ocorrências têm na tradução?; (v) Que padrões de tradução podem ser identificados em cada um dos contextos?

Pode-se dizer que os problemas da tradução não se reduzem simplesmente a problemas de vocabulário trata-se de um complexo processo de transposição interlíngual em que estão envolvidos aspectos culturais e linguísticos. Assim, glossários ou dicionários sem exemplos reais e contextualizados da língua não resolvem todos os problemas tradutórios.

Nessa perspectiva, os avanços tecnológicos podem servir como ferramenta de apoio para a tradução ou para o ensino da tradução. A familiaridade com essas ferramentas ajuda profissionais, pesquisadores e alunos a explorar ao máximo as possibilidades tradutórias de suas áreas.

O estudo da lexicografia é um instrumento valioso no estudo das línguas, sendo que, a elaboração de dicionários existe desde a Antiguidade, pois o homem cedo percebeu a necessidade de ter ferramentas auxiliares à comunicação.

O fato de existir mais de um idioma faz com que haja a necessidade de se compreenderem suas diferenças linguísticas. Segundo Philippe Humblé (2001), todos os dicionários ou glossários que conhecemos até o Renascimento eram bilíngues no sentido *interlingual*<sup>1</sup>

---

<sup>1</sup> Intralingual, ou paráfrase – “que é reescrever um texto a partir dos signos de uma mesma língua” disponível em 18 de março de 2010 em <http://fibonacci55.com/blogs/ernesto/2009/11/11/traducao-intersemiotica-rapidamente/>

ou *intralingual*<sup>2</sup>. Esses glossários ofereciam traduções de palavras da língua A para a língua B ou da língua A para a língua A com palavras mais simples, como é o caso dos glossários gregos de Homero, que traduziam palavras do grego antigo para um grego menos antigo. Humblé (2001) afirma que, somente em 1612, a *Accademia della Crusca* produziu um dicionário monolíngue, provavelmente para padronizar uma língua em uma época em que cada um escrevia da forma como achava que deveria ser escrita. Com o passar dos anos, os dicionários foram tomando formas diferentes, como se depreende das palavras de Humblé ao afirmar que

A partir do momento em que a *produção* de uma língua estrangeira começou a ser importante, os lexicógrafos começaram a perceber que as duas atividades, *produção* e *compreensão*, seguiam estratégias diferentes e requeriam uma organização específica em nível de organização do dicionário (HUMBLÉ, 2001).

Quanto ao processo de compreensão, é importante que um dicionário bilíngue apresente várias acepções de tradução para um termo e, principalmente, que se preocupe com as palavras mais raras, pois provavelmente o tradutor vai procurar por estas no dicionário, ao contrário do aprendiz iniciante. Já quanto ao processo de produção de um texto o percurso se torna mais difícil, pois não há texto contextualizado, como é o caso na compreensão do texto. Nesse sentido, o tradutor precisa de um dicionário com muitos exemplos, para que possa construir frases que soem naturalmente dentro de uma padronização aceita por falantes nativos. Béjoint (1981, p. 210) afirma que “Em geral, o melhor dicionário para a produção é aquele que dá as informações mais detalhadas sobre sintaxe e colocação [...]” Tanto a sintaxe como a colocação estudam os processos combinatórios das frases, ou seja, suas co-ocorrências dentro de determinado contexto. No *Compara* facilmente encontram-se as ocorrências de: *hearing about*, *on hearing this* ou *hearing something*. Por isso, a importância de exemplos de frases nos dicionários, mas esses exemplos devem ser reais da língua e não inventados. Segundo Humblé

---

<sup>2</sup> Interlingual – “a tradução propriamente dita, que seria passar de uma língua para a outra” disponível em 18 de março de 2010 em <<http://fibonacci55.com/blogs/ernesto/2009/11/11/traducao-intersemiotica-rapidamente/>>

(2001) quando surge a consciência desses fatos na linguística, percebe-se a necessidade do uso de *corpora* como ferramenta de apoio.

Com a necessidade de exemplos reais da língua, mais o surgimento dos microcomputadores e também com o aumento da capacidade e velocidade no armazenamento e processamento dos dados, a situação muda consideravelmente, possibilitando aos especialistas desempenharem uma função importante no desenvolvimento de áreas como a linguística, a lexicografia e a tradução. Com o avanço tecnológico, pode-se atualmente fazer uma análise de *corpora paralelos*<sup>3</sup> mais rápida e eficaz, dando vantagens ao pesquisador quanto à obtenção de dados de uma língua, devido à rapidez e agilidade dos computadores. Segundo Sinclair (1991), “Corpus is a collection of naturally occurring language text, chosen to characterize a state or a variety of a language”<sup>4</sup>. Entende-se então, que *corpora* são textos autênticos eletronicamente armazenados, representativos de uma determinada língua ou área a ser estudada.

Podemos identificar dois períodos importantes desde o surgimento da linguística de *corpus*: o antes e o depois da era tecnológica. Antes do computador, já se fazia uso de *corpus*, quando, na Grécia Antiga, foi criado o *Corpus* Helenístico. Segundo Berber Sardinha (2000), na Antiguidade e na Idade Média produziam-se *corpora* de citações da Bíblia. Em grande parte do século XX, usavam-se *corpora* para fazer a descrição da linguagem, porém os *corpora* dessa época eram coletados e depois analisados manualmente, o que dificultava a realização de pesquisas desse tipo.

Nessa época, os estudos com base em *corpus* manuais, chamados empíricos<sup>5</sup>, receberam forte crítica devido à necessidade de se coletarem dados dos quais não se tinha uma confiabilidade em analisar manualmente grandes quantidades de dados linguísticos, contudo os estudos baseados em *corpora* não pararam. Vários pesquisadores, assim como Firth (1957) e os neo-firthianos acreditavam na descrição da linguagem por meio de dados reais. O *corpus* SEU (Survey of English Usage) foi compilado e etiquetado manualmente a partir de 1953 e influenciou a

---

<sup>3</sup> Segundo Baker (1995) um corpus paralelo é uma coleção de textos originais (língua fonte) e suas versões traduzidas.

<sup>4</sup> Corpus é uma coleção de ocorrências de textos naturais da língua, selecionados para caracterizar um estado ou variedade de uma língua. (Tradução da autora).

<sup>5</sup> O estudo empírico se baseia na experiência ou observação de dados autênticos reunidos em um banco de dados. Disponível em 18 de março de 2010 em: <http://pt.wikipedia.org/w/index.php?title=Especial%3ABusca&redirs=0&search=empirico+linguistica&fulltext=Search&ns0=1&searchengine=select=mediawiki>.

criação de *corpora* eletrônicos que serviram para o desenvolvimento de etiquetadores computadorizados contemporâneos, segundo Berber Sardinha (2000).

Com o advento do computador e também com a queda de prestígio das pesquisas racionalistas<sup>6</sup>, o cenário começou a mudar. O lançamento do *Brown Standard Corpus*, em 1964, com 1 milhão de palavras, é considerado como o fato propulsor do desenvolvimento da Linguística de Corpus. O *corpus* Brown foi o primeiro dos *corpora* eletrônicos. A popularização dos estudos com *corpora* ocorreu na década de 1980, com o aparecimento dos computadores pessoais.

Ao final dos anos de 1990, a Linguística de Corpus exerceu grande influência em várias áreas da Linguística Aplicada, assim como no ensino de línguas, na tradução e no ensino de tradução, na análise do discurso, na lexicografia, na elaboração de material didático, entre outros, fornecendo suportes teóricos e metodológicos, que possibilitem o estudo de vários aspectos da linguagem.

Tendo destaque como ferramenta de apoio, os *corpora* eletrônicos são hoje considerados por muitos estudiosos essenciais para o profissional da tradução, tanto nas traduções literárias como terminológicas. Berber Sardinha (2003, p. 70) afirma que

A organização da linguagem é muito mais complexa do que o suposto, ficando claro que a utilização de *corpora*, de certa forma, deixou de ser uma opção para investigações na área.

A utilização de *corpora* por tradutores funciona como um sistema de auxílio tanto para a pesquisa quanto para a prática profissional por possibilitar a exploração de diversos aspectos linguísticos e tradutológicos.

## ORGANIZAÇÃO DO TRABALHO

A divisão deste trabalho ocorrerá da seguinte maneira; no primeiro capítulo tratará da: (i) Revisão da Literatura, na qual serão apre-

---

<sup>6</sup> O estudo racionalista baseia-se no conhecimento estabelecido a priori, através da lógica dedutiva, da introspecção e dos processos cognitivos. Disponível em 18 de março de 2010 em: <<http://www.fafich.ufmg.br/cogvila/dischistoria/gauer1.pdf>>

sentados alguns autores importantes da área de pesquisa; (ii) definição de *corpus*; (iii) finalidade de um *corpus*; (iv) disciplina de estudos da tradução; (v) uso de *corpora* em estudos da tradução; (vi) tipos de *corpora*; e (vii) Linguística de Corpus e a importância da terminologia na tradução. O segundo capítulo abordará a metodologia para a realização das atividades de pesquisa, o desenho do *corpus EuroParl* e do *Compara*, os procedimentos de coleta da amostra do item lexical *hearing* e o procedimento para o processamento do *corpus*. E no terceiro capítulo, será apresentada a análise dos dados e resultados encontrados.

## CAPÍTULO 1

### 1 REVISÃO DA LITERATURA

A revisão da literatura que fornece embasamento teórico para o presente trabalho está organizada da seguinte forma: (i) Definição de *corpus*; (ii) A finalidade de um *corpus*; (iii) Estudos da Tradução; (iv) *Corpora* em Estudos da Tradução; (v) Linguística de Corpus; e (vi) Importância da terminologia na tradução.

#### 1.1 Definição de *Corpus*

Primeiramente a palavra *corpus* era atribuída a qualquer conjunto de textos. Hoje um *corpus* é uma coleção de textos, que devem ser autênticos<sup>7</sup> e não inventados, capturados e armazenados de forma eletrônica que são representativos dentro de uma língua ou variedade linguística. Baker (1995) diz ainda que um *corpus* deve conter vários textos, de diversas fontes e escritos por diversos autores, para atender a uma necessidade específica.

Corpus now means primarily a collection of texts held in machine-readable form and capable of being analyzed automatically or semi-automatically in a variety of ways: a corpus may include a large number of texts from a variety of sources, by many writers and speakers and a multiple of topics. What is important is that it is put together for a particular purpose and according to explicit design criteria in order to ensure that it is representative of the given area or sample of language it aims to account for (BAKER, 1995:225)<sup>8</sup>.

---

<sup>7</sup> Textos autênticos são os textos falados ou escritos em situação real da língua, e não aqueles produzidos apenas para representar a língua (conclusão da autora).

<sup>8</sup> Corpus, hoje, significa essencialmente uma coleção de textos agrupados em forma legível eletronicamente e capaz de ser analisado automaticamente ou semi - automaticamente em diversas maneiras: um corpus deve incluir uma grande quantidade de textos de diversas fontes, por diversos autores falantes sobre vários tópicos. O mais importante é, que o corpus seja compilado com o objetivo de atender uma necessidade específica de acordo com critérios específicos e que seja uma amostra representativa da língua dentro da área que se queira estudar. (Tradução da autora).

O objetivo do corpus é facilitar a pesquisa acadêmica ou profissional sobre o léxico e a gramática, ou ainda ajudar nas investigações históricas ou culturais da língua. Os corpora são mais utilizados entre os linguistas interessados em vários aspectos da língua. No entanto, os corpora são úteis para não linguistas também. Pesquisadores de literatura podem utilizar-se do corpus para análises linguísticas de aspectos de um determinado período de tempo na literatura. Professores de línguas estrangeiras ou não podem se beneficiar com o uso de corpus para investigações estatísticas quanto à padronização da língua, entre outros estudos.

## 1.2 Qual a finalidade de um corpus?

Existe uma variedade de análises linguísticas que se pode fazer com um *corpus* e várias áreas que podem beneficiar-se do uso de *corpus*. Entre elas destacam-se:

- Verificar a frequência das palavras mais comuns dentro de uma área específica.
- Analisar a frequência das classes gramaticais.
- Reconher e detalhar as lexias complexas e compostas<sup>9</sup>.
- Observar e analisar os padrões de uma determinada língua por meio das colocações (a colocação indica uma combinação provável mais aceita pelos falantes nativos, como, por exemplo: andar de ônibus, andar a cavalo, andar triste, o quinto andar do prédio, etc).
- Analisar as estruturas idiomáticas da língua a ser aprendida e comparar com a língua materna.
- Analisar um *corpus* de aprendiz, para descobrir quais os erros mais comuns cometidos por estudantes que falam o português como língua materna e comparar com os erros de falantes nativos de outras línguas, assim como um falante nativo de francês ou japonês que fala o inglês também como segunda língua.
- Utilizar o *corpus* de aprendiz para elaborar material didático.
- Servir como base para o estudo da lexicografia pedagógica.

---

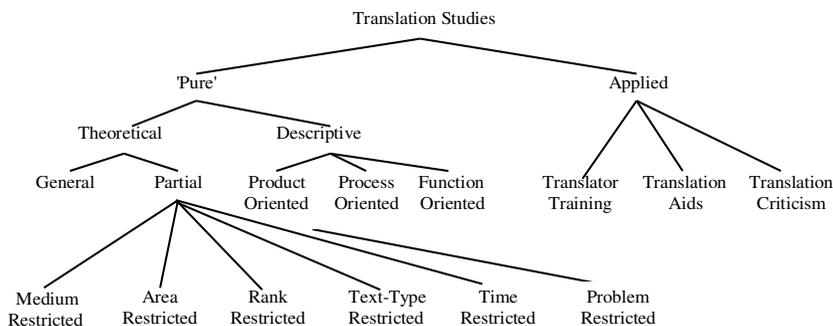
<sup>9</sup> As lexias compostas seriam geradas por duas palavras (justapostas ou hifenizadas), gerando uma terceira palavra e um terceiro sentido, porém ainda guardando uma relação de significação com os dois sentidos originais. Ex.: *guarda-chuva*. As lexias complexas trabalham no nível frasal, onde podemos ou não recuperar o sentido original de cada lexia simples, mas o sentido final é independente delas. Ex.: *certificado de depósito bancário* (CDB).

- Produzir dicionários bilíngues ou multilíngues.
- Ou ainda pode servir como base de dados para os tradutores.

Enfim, podemos fazer várias análises linguísticas e em diversos campos do saber, como na terminologia, na lexicografia, na tradução, no ensino de tradução, na construção de material didático, no ensino de línguas, por alunos iniciantes ou avançados, por alunos presenciais ou a distância. Contudo, ao fazer uma coleta ou seleção de um *corpus* é necessário estabelecer alguns critérios. Devemos nos perguntar para qual finalidade se quer aplicar o *corpus*, para quem se destina o *corpus*, quais línguas serão trabalhadas, qual o tamanho do *corpus*, ou seja, precisamos desenhar o *corpus* e o propósito.

### 1.3 Estudos da Tradução

Holmes iniciou um movimento político acadêmico através do qual queria criar um espaço privilegiado para o desenvolvimento de uma reflexão sobre a teoria e a prática da tradução. Os Estudos da Tradução, como disciplina acadêmica, só passou a existir oficialmente na segunda metade do século XX graças a Holmes, que apresentou pela primeira vez um artigo com o nome de *The name and nature of translation studies* no Setor de Tradução do Terceiro Congresso Internacional de Linguística Aplicada, realizado em Copenhague em 1972, mas somente em 1988 é que foi oficializado como disciplina independente. Holmes (1972, 1988) fez um mapeamento sugerindo o que a disciplina de Estudos da Tradução cobriria, como apresentado na Figura 1:



**Figure 1 - Holmes's map of translation (from Toury 1995:10)**

O Mapa de Holmes *Translation Studies* representado na Figura 1 por Toury (1995) serve como ponto de partida para a disciplina de Estudos da Tradução. O autor propõe não somente o nome para a disciplina de Estudos da Tradução, como também nomeia alguns ramos de pesquisa de estudo dividindo em estudos aplicados e puros.

Segundo Munday (2001), os ramos dos estudos da tradução puros têm dois principais objetivos: (i) *descritivos*, que têm como objetivo descrever o produto, ou seja, procuram descrever os fenômenos da tradução (chamados de Descriptive Translation Studies – DTS) e os (ii) *teóricos*, que tentam estabelecer princípios gerais de significados para que estes fenômenos possam ser explicados e previstos (chamados de Translation Theory – TTh).

Dentro dos *estudos descritivos*, Holmes aponta três outras áreas de estudo: (i) *produto*, que descreve a tradução já existente; (ii) *processo*, que procura entender o processo da tradução na mente do tradutor; e (iii) *função*, que analisa a função da tradução no contexto sócio cultural, ou seja, a função da tradução na cultura receptora, que estuda mais o contexto do que propriamente o texto.

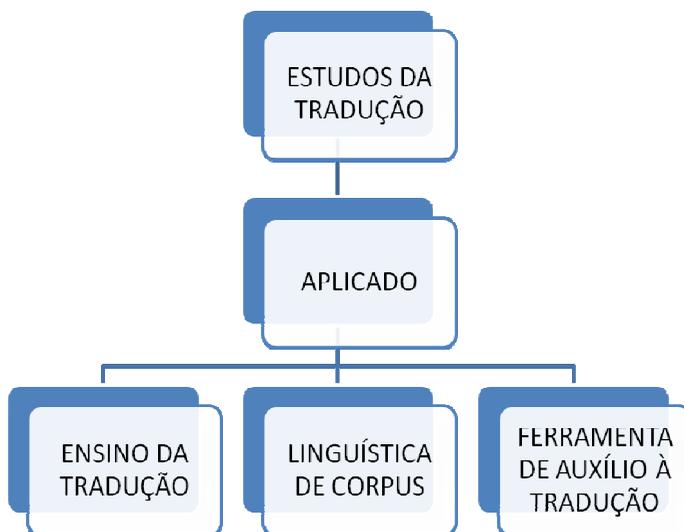
Dentro do ramo dos *estudos teóricos*, o autor dividiu-os em duas categorias: estudos teóricos gerais e estudos teóricos parciais (ou específicos). Os *estudos teóricos gerais* servem para explicar ou prever todos os fenômenos que acontecem dentro do terreno da tradução, enquanto que os *estudos teóricos parciais* analisam uma categoria mais restrita das traduções.

Os *estudos teóricos parciais* são divididos em seis: (i) restrito ao problema: analisam problemas de traduções – questões como variáveis e invariáveis e a equivalência; (ii) restrito à área: estudo relacionado a restrição de língua e cultura; (iii) restrito ao nível: teorias que lidam com o discurso ou textos como um todo; (iv) restrito ao tipo de texto: lidam com problemas de gêneros ou tipos específicos de tradução; (v) restrito ao tempo: analisam traduções de textos contemporâneos e tradução de textos de um período antigo; (vi) restrito ao meio: são traduções humanas/traduções por máquina.

Os *estudos da tradução aplicados* (TS) seguem quatro principais linhas de pesquisa: (i) ensino da tradução, que se preocupa com a formação do tradutor, incluindo, métodos de ensino, técnicas de avaliação e planos de ensino; (ii) ferramentas de auxílio à tradução, assim como dicionários, gramáticas e tecnologias; (iii) crítica da tradução, que se preocupa com a avaliação, crítica ou análise das traduções incluindo revisões de traduções publicadas e traduções de alunos; e (iv) política da

tradução, que determina o que precisa ser traduzido em uma determinada posição cultural e a posição cultural econômica do tradutor.

Segundo Pagano & Vasconcellos (2003), a sugestão de mapeamento da disciplina de Estudos da Tradução de Holmes foi feita na década de 1970, quando muitos avanços em diferentes áreas do conhecimento ainda não tinham acontecido, principalmente quando o avanço tecnológico e as ferramentas de apoio que possibilitam novas pesquisas não haviam sido descobertos ainda. Hoje, há outros ramos de estudo dentro dos Estudos da Tradução, assim como, os Estudos da Tradução com base em *corpus* e o estudo da multimídia (legendagem e estudo das traduções), entre outros que não se encontram no mapeamento de Holmes. Baseada no mapeamento de Holmes (1972) e nas ideias de Pagano & Vasconcellos (2003) acima citadas a autora sugere um mapeamento da presente pesquisa:



**Figura 2 – Mapeamento sugerido pela autora (Marian, 2010)**

A presente pesquisa está inserida dentro da disciplina de estudos da tradução, é um estudo Aplicado da língua, pois apresentam dados reais da língua através do qual se pode analisar características linguísticas. O estudo faz parte linguística de Corpus, que estuda a língua através

dos dados e serve como Ferramenta de Auxílio à Tradução e Ensino da Tradução.

#### 1.4 Corpora em Estudos da Tradução

A utilização dos *corpora* para análises textuais permite a realização de análises necessárias para o estudo da língua, como por exemplo: (i) análises estatísticas: verificar o número de ocorrências de um determinado elemento linguístico, para saber estatisticamente quais os padrões são mais usados por um falante nativo; (ii) confrontar regras linguísticas de uma determinada área como, por exemplo, o uso de *z* e *s* no português; (iii) codificar itens lexicais, por exemplo, *guide* (verb), *guide* (noun) ou outros símbolos para uma visualização textual mais compacta e eficiente; (v) investigar quantitativamente aspectos que caracterizam os textos escritos por diferentes culturas e questionar grandes quantidades de dados da qual nenhum ser humano em vida poderia gerenciar de forma tão rápida e eficaz sem a ajuda dos meios eletrônicos de hoje. Para Tymoczko (1998, p.1):

[...] corpus translations studies change in a qualitative as well as a quantitative way both the content and the methods of the discipline of translation studies, in a way that fits with the modes of the information age.<sup>10</sup>

O estudo da tradução a partir de dados fornecidos por corpora mudam a visão da língua tanto de forma quantitativa quanto qualitativa. De forma quantitativa, pois o corpus apresenta um grande número de exemplos onde podemos visualizar um determinado léxico contextualizado. E qualitativo, pois os exemplos são exertos de composição ou fala da vida real de falantes nativos.

Para a realização de uma pesquisa de tradução baseada em *corpus* faz-se necessário conhecer e classificar os tipos de *corpus* a serem estudados. Baker (1995) e Tagnin (2005) distinguem três tipos de *corpora*: *corpora* paralelos, *corpora* comparáveis e *corpora* multilíngues, os quais serão explanados nos próximos tópicos.

---

<sup>10</sup> Os estudos da tradução de corpus mudam tanto de uma forma qualitativa quanto qualitativa o conteúdo e métodos da disciplina de estudos da tradução, de maneira a se encaixar nos moldes da era da informação. (Tradução da autora).

### 1.4.1 Corpora Paralelos

Baker (1995) define *corpus* paralelo como uma coleção de textos originais (língua fonte) e suas versões traduzidas. Para Baker, os *corpora* paralelos ajudam o tradutor a lidar com os problemas de tradução na prática, além de fornecerem modelos realistas para tradutores iniciantes.

They allow us to stabilish, objectively, how translators overcome difficulties of translation in practice, and to use this evidence to provide realistic models for trainee translators. They also have an important role to play in exploring norms of translating in specific social-cultural and historical contexts<sup>11</sup> (BAKER, 1995, p.231)

### 1.4.2 Corpora Comparáveis

Para Tagnin (2005), os *corpora* comparáveis<sup>12</sup> são *corpora* bilíngues constituídos de textos sobre os mesmos assuntos ou áreas, gênero, tipologia, extensão e data de publicação. Os *corpora* comparáveis podem ajudar o tradutor na investigação de termos equivalentes na outra língua, já que os *corpora* paralelos são restritos e de pouco acesso, principalmente quando se tratam áreas técnicas.

### 1.4.3 Corpora Multilíngues

Para Baker (1995), os *corpora* multilíngues são conjuntos de dois ou mais *corpora* monolíngues em diferentes línguas, que nos permitem estudar itens e elementos linguísticos em seu ambiente natural, em vez de como são utilizados em textos traduzidos, e, portanto têm um papel importante a desenvolver no treinamento de tradutores, escrita de materiais e melhorias de sistemas de tradução automática.

---

<sup>11</sup> Eles nos permitem estabelecer, objetivamente, como tradutores superam dificuldades de tradução na prática e a usar esta evidência para fornecer modelos realísticos para tradutores aprendizes. Eles também têm um papel importante em explorar normas de traduzir em contextos sócio-culturais e históricos específicos. (Tradução da autora).

<sup>12</sup> Corpora Comparáveis: Disponível em maio 2010 em: <[HTTP://www.fflch.usp.br/](http://www.fflch.usp.br/)> A identificação de equivalentes tradutórios em corpora comparáveis (Stella Tagnin).

### 1.4.4 Classificação do Corpus

Baker (1995) também propõe critérios para a classificação dos *corpora* nos ramos descritivos e aplicados dos Estudos da Tradução baseados em *corpora*, o desenho do *corpus* é um dos primeiros estágios para um estudo baseado em *corpus*, pois se entende como o planejamento do *corpus* de estudo.

(i) General language vs. restricted domain; (ii) written vs. spoken language; (iii) synchronic vs. diachronic; (iv) typicality in terms of range of sources (writers/speakers) and genres (e.g. newspaper editorials, radio interviews, fiction, journal articles, court hearings); geographical limits, e.g. British vs. American English; monolingual vs. bilingual of multilingual<sup>13</sup> (BAKER, 1995, p.229).

Para fazer o planejamento do *corpus* é necessário conhecer sua tipologia, tal como um *corpus* pode ser composto por textos em uma única língua, denominado monolíngue; em duas línguas, denominado bilíngue; ou em mais línguas, denominado multilíngue. Segundo Baker (1995), para os *Estudos da Tradução* assume-se que um *corpus* não pode ser apenas monolíngue em uma única língua, deve, pelo menos, envolver duas ou mais línguas quando se trata de um estudo contrastivo.

A tabela abaixo foi desenhada baseando-se em alguns critérios citados por Baker (1995), Sardinha (2000;2004) e Fernandes (2006):

<i>PROPOSTA DE TIPOLOGIA DO CORPUS</i>	
<b>Domínio:</b>	Domínio pode ser <i>geral</i> (revistas, jornais, literários, etc.) ou <i>específico</i> (jurídico, por exemplo).
<b>Modo:</b>	Podem ser compostos por textos <i>escritos</i> ou <i>falados</i> .
<b>Temporalidade:</b>	O <i>corpus</i> pode ser <i>sincrônico</i> (quando é um só período) ou <i>diacrônico</i> (quando se refere a vários

<sup>13</sup> (i) língua geral versus domínio restrito; (ii) língua escrita versus falada; (iii) sincrônico versus diacrônico; (iv) tipicidade em termos de variedade de fontes (escritas/faladas) e gêneros (ex. jornais, rádios, audiências, artigos de jornais); limites geográficos ex., inglês britânico versus americano; monolíngue versus bilíngue. (Tradução da autora).

	períodos de tempo)
<b>Classificação</b>	<i>Paralelos</i> (textos fontes e suas traduções), ou <i>Comparáveis</i> (textos originais de uma área em L1 e textos originais da mesma área em L2) ou <i>Multi-lingues</i> (textos monolíngues em diferentes línguas).
<b>Número de línguas</b>	<i>Monolíngue</i> (uma língua), <i>bilíngue</i> (duas línguas) ou <i>multilíngue</i> (mais de duas línguas).
<b>Direcionalidade</b>	<i>Unidirecional</i> (uma direção – do português para o inglês somente), <i>Bidirecional</i> (duas direções – do português para o inglês e vice-versa) ou <i>Multidirecional</i> (com mais de duas direções – do inglês para o português – do português para o Francês – do Francês para o alemão, etc.).
<b>Tamanho</b>	Quanto ao tamanho do corpus (veja Tabela 1 – SARDINHA, 2000)

### Quadro 1: Tipologia do Corpus (síntese da autora)

Quanto ao último tópico do Quadro acima - Tamanho do Corpus - Berber Sardinha (2000), denomina de representatividade de um *corpus*. Mas qual seria o tamanho de uma amostra representativa dentro de uma língua/área a ser estudada? Não há um tamanho exato do que é um *corpus* representativo dentro de uma determinada área de estudo, depende do objetivo do estudo. O que Berber Sardinha (2000) sugere é que o *corpus* deve ser o maior possível dentro da área específica, pois a linguagem é um sistema probabilístico<sup>14</sup>, então quanto maior o *corpus* maior será a probabilidade de algumas palavras de frequências raras aparecerem. Dessa forma, quando se fala de representatividade, refere-se à extensão do *corpus* em número de palavras e textos. Berber Sardinha (ibid.) propõe uma sugestão de *classificação de corpus* baseada em

<sup>14</sup> Segundo Beber Sardinha (2002), a linguagem é um sistema probabilístico porque envolve escolhas, quando produzimos um texto temos que escolher qual item lexical vamos usar com determinada palavra, essas escolhas estão ligadas a probabilidades de usar um ou outro item lexical.

observações de *corpora* utilizados durante quatro anos em conferências de Linguística de Corpus:

**Tabela 1: Tamanho do Corpus**

Menos de 80 mil	Pequeno
80 a 250 mil	Pequeno-médio
250 mil a 1 milhão	Médio
1 milhão a 10 milhões	Médio-grande
10 milhões ou mais	Grande

Fonte: Berber Sardinha (2000)

De acordo com esse critério, pode-se afirmar que o presente *corpus* de estudo, o *EuroParl* é um *corpus* de tamanho grande, portanto podemos dizer que ele é ou deveria ser representativo dentro da área legislativa. E, na mesma perspectiva, o *corpus Compara* pode ser considerado de médio-grande porte, o que também é considerado como um *corpus* representativo dentro da área literária.

## 1.5 Linguística de Corpus

Tony Berber Sardinha (2002), da Pontifícia Universidade Católica de São Paulo, é um pesquisador que praticamente introduziu a Linguística de Corpus no Brasil. O autor afirma que a Linguística de Corpus baseia-se na coleta e exploração de *corpora*, ou conjunto de dados linguísticos textuais que são coletados criteriosamente com o propósito de servirem para a pesquisa de uma língua ou variedade linguística e dedica-se à exploração da linguagem por meio de evidências empíricas, extraídas por computador. Este permite que grandes bancos de dados sejam analisados de forma mais confiável e rápida. Esse contexto tecnológico permite a análise de um item lexical e suas colocações.

A Linguística de Corpus caracteriza-se pela coleta e análise de dados linguísticos processados por ferramentas eletrônicas. O computador desempenha um papel importante para os estudos na área, pois as ferramentas computacionais são geralmente utilizadas para reorganização e extração de informações do *corpus* para observação e interpretação de dados, fornecendo novas perspectivas para a análise linguística.

A Linguística de Corpus faz uso de uma abordagem empirista contrária à abordagem racionalista, do ponto de vista linguístico, e tem

como eixo central a noção de linguagem como sistema probabilístico. De acordo com essa noção, os traços linguísticos não ocorrem de forma aleatória, sendo possível evidenciar e quantificar padrões. É comum nesta área afirmar que a linguagem é padronizada, isto é, existe uma correlação entre os traços linguísticos e os contextos situacionais de uso da linguagem.

Na Linguística de Corpus, a padronização se evidencia por colocações, coligações e prosódia semântica que se repetem de forma significativa. De acordo com essa visão, a descrição de uma palavra e seus padrões não pode ser considerada uma questão apenas lexical ou gramatical, já que os sentidos e a padronização estão conectados. Por esse motivo, a Linguística de Corpus se opõe à divisão tradicional entre léxico e gramática e considera que há um nível do sistema linguístico que engloba tanto o vocabulário quanto a gramática, o que denomina léxico-gramática.

Segundo Berber Sardinha (2002), se a produção de textos envolve escolha, estas estão obviamente ligadas a uma probabilidade. Então, de acordo com isso, cada escolha tem uma probabilidade maior ou menor de ser selecionada pelo falante. Existe uma relação de expectativa que deve ser mantida pelo falante de uma língua para seguir o padrão de naturalidade.

A escolha de cada item lexical implica na redução das escolhas dos itens lexicais e das categorias gramaticais que podem segui-lo. A escolha de uma classe gramatical reduz as escolhas possíveis de classes gramaticais e itens lexicais que podem seguir-se a ela (BERBER SARDINHA. 2000, p.49).

De acordo com Berber Sardinha (ibid), as escolhas das palavras reduzem as opções gramaticais que podem seguir a mesma palavra. Cada item lexical tem algumas opções que podemos usar para formar frases, que são chamadas de padronização da língua.

It's known that what a word means often depends on its association with certain collocates, with what comes after/or before it, and these choices come from cultural, social and historical use of language<sup>15</sup> (SINCLAIR, 2001).

---

<sup>15</sup> Sabe-se que o significado de uma palavra depende da sua associação com certo colocado, com o que vem antes ou depois da palavra, e estas escolhas são culturais, sociais e históricas da língua. (Tradução da autora).

Essa regularidade (padrão) da língua depende de fatores culturais, sociais e históricos da língua. Portanto, conhecer os itens lexicais e seus colocados pode implicar ter fluência ou não de uma língua, pois essa padronização soa familiar para os falantes nativos.

Berber Sardinha (2002) também distingue três tipos de padrões léxico-gramaticais: (i) a *colocação*<sup>16</sup> – refere-se à associação entre os itens lexicais. Sinclair (ibid), define colocação como “the occurrence of two or more words within a short space of each other in a text”<sup>17</sup> (p. 170). Na colocação, aparecem co-ocorrências que se repetem frequentemente ou que são estatisticamente consideráveis, por exemplo: *happy birthday, credit card, make a joke, take care, Merry Christmas*, entre outros exemplos que poderíamos citar. Tagnin (2005), ao citar *Merry Christmas* em seu livro *O jeito que a gente diz* afirma que, embora se use *Happy New Year*, não é comum ouvir dizer *Happy Christmas*, pois existe uma convencionalidade quanto ao uso de *Merry Christmas*. Para o termo convencional, Tagnin (2005, p.15) o define como tudo que é convencional, ou seja, tudo que é comum, de praxe, consolidado pelo uso ou pela prática. *Merry Christmas* é uma convenção social, um costume estabelecido pela sociedade, que acabou virando regra, estando ou não de acordo com as regras gramaticais. Tagnin (2005, p.30) define coligação por “combinação consagrada de elementos linguísticos em que o colocado é uma palavra gramatical, por exemplo: look at, mad about...”.

A colocação é a evidência de que as palavras não se combinam por acaso. Existem algumas regras gramaticais, mas muitas colocações fogem a estas regras e ignorar essas combinações pode causar problemas de compreensão para o leitor nativo. Para Sinclair (2001) as palavras não ocorrem aleatoriamente em um texto, pois temos um grande número de expressões pré-construídas que se constituem em escolhas que fizemos quando elaboramos um diálogo ou texto:

---

<sup>16</sup> Segundo Tagnin (2005) as coligações e colocações se referem a uma combinação gramatical ou lexical, respectivamente. Tagnin (2005, p.30) afirma: “**Coligação**: combinação consagrada de elementos linguísticos em que o colocado é uma palavra gramatical, por exemplo, em inglês, *look at, mad about*, em português, *obedecer a, cumpridor de*. As *coligações* podem ser classificadas como: *coligações de regência* (verbos, substantivos, adjetivos ou advérbios), *verbos frasais* (esta classe não existe em português, são os chamados *phrasal verbs* em inglês, formados por um verbo seguido de uma partícula adverbial) ou *coligações prepositivas*.”

**Colocação**: combinação lexical consagrada de duas ou mais palavras de conteúdo: por exemplo, dizemos *red cabbage* em inglês, mas *repolho roxo* em português; *rocking chair* em inglês, mas *cadeira de balanço* em português e não *cadeira balançante*”.

<sup>17</sup> A ocorrência de duas ou mais palavras dentro de um pequeno espaço em um texto. (Tradução da autora).

It is clear that words do not occur at random in a text, and that the open-choice principle does not provide for substantial enough restraints on consecutive choices...The principle of idiom is that a language user has available to him or her a large number of semi-preconstructed phrases that constitute single choices...<sup>18</sup> (SINCLAIR, 2001, p.110).

A (ii) *coligação* é definida por Berber Sardinha (ibid.), como a associação entre os itens lexicais e gramaticais. E de acordo com Sinclair (2001), o sentido das palavras no contexto altera a escolha do léxico usado para o restante da frase: "... the different meanings of a word have very different frequencies"<sup>19</sup> (p. 56). A análise das coligações ajuda o tradutor a perceber que a mesma palavra poderá ter coligações diferentes dependendo do significado que a palavra vier a ter. Por exemplo, a palavra *start* é mais usada com frases *ing* e sintagmas nominais, enquanto *begin* é mais usado com *to*.

E o último (iii), a *prosódia semântica*, segundo Berber Sardinha (ibid) é uma associação entre os itens lexicais e a conotação de campos semânticos (negativa, positiva ou neutra). Assim como na fala, a prosódia prepara o interlocutor para tipos de sons que virão a seguir, e por sua vez, a prosódia semântica prepara o leitor para o conteúdo semântico. Por exemplo, a palavra *cause*, citada por Sinclair (1991)<sup>20</sup>, tem um conteúdo semântico negativo, pois se associa com palavras de valor negativo, como *cause damage/problems/of death/suffering/fear/an accident*. Essa ocorrência pode justificar a fluência ou não fluência de uma língua, já que os traços linguísticos e o léxico criam *relações de expectativas* que devem ser mantidas pelo falante de uma língua mantendo o padrão de naturalidade. No entanto, esse estudo não pretende abordar a prosódia semântica do item de pesquisa.

A descoberta de que uma palavra está padronizada tanto lexicalmente quanto gramaticalmente baseia-se em estatísticas por meio das concordâncias. Estas são listas de orações extraídas de um *corpus*, com-

---

<sup>18</sup> Está claro que as palavras não ocorrem de forma aleatória em um texto, e que o princípio de escolha aberta não fornece significado suficiente para escolhas sucessivas. O princípio de um idioma é que o usuário tem disponível um grande número de frases pré-construídas que constituem únicas escolhas. (Tradução da autora).

<sup>19</sup> ...significados diferentes de uma palavra têm frequências diferentes. (Tradução da autora).

<sup>20</sup> Sinclair, John M. (1991). "Words and phrases". *Corpus, Concordance, Collocation*. Oxford: OUP. pp. 70-75.

pletas ou não, nas quais uma ou mais palavras aparecem centralizadas. A palavra central recebe o nome de *nódulo* e o que fica à direita e à esquerda do texto chama-se de co-texto. As concordâncias permitem a análise de um grande número de dados dos quais é possível calcular a frequência de co-ocorrências de palavras.

## 1.6 A Importância da Terminologia na Tradução

Hoje, com o mundo dos negócios cada vez mais globalizado e a tecnologia mais avançada, as pesquisas em dicionários convencionais vêm se tornando uma raridade pelos pesquisadores que buscam por dicionários *on-line* para a obtenção de respostas mais rápidas e confiáveis.

Na *web*, temos algumas ferramentas de apoio ao tradutor. Existem *corpora* paralelos de tradução especializada como: (i) o *EuroParl*<sup>21</sup>, Constituição e o *JRC* (*corpora* da área legislativa) que estão alinhados em inglês/português disponíveis em *Natools Corpora Query Interface*<sup>22</sup>; (ii) *Compara*<sup>23</sup> (*corpora* literários), alinhado em português/inglês, é um projeto de investigação científica sem fins lucrativos (parte integrante do projeto Linguateca); (iii) o *Turigal*<sup>24</sup>, (*corpora* da área de turismo), alinhado em português/inglês, um projeto *Cluvi* (*Corpus* Linguístico da Universidade de Vigo); (iv) *CorTrad*<sup>25</sup> (culinária e divulgação científica) alinhados em português/inglês, é um *corpus* multilíngue para o ensino e tradução (projeto do *Comet*) entre outros que existem disponíveis *on-line*.

Os *corpora* podem ser de grande ajuda na confirmação de decisões, reforçando modelos de padrões da linguagem ou quanto ao uso de colocações (palavras que co-ocorrem normalmente ao redor de um item lexical). Consultar um *corpus* paralelo na língua alvo, seja para pesquisar terminologias ou buscar por frases idiomáticas é algo que vem se tornando familiar, ou pelo menos deveria ser para os tradutores profissionais, pois lhes permite explorar grande quantidade de textos de forma rápida e precisa. Quando decidimos por um equivalente de um termo ou frase para a tradução, os *corpora* podem oferecer ampla ajuda na decisão de rejeição ou aceitação de um item lexical de valor aproximado.

No capítulo a seguir, será abordada a metodologia da pesquisa.

---

<sup>21</sup> <http://www.statmt.org/europarl>

<sup>22</sup> <http://linguateca.di.uminho.pt/nat/nat.pl>

<sup>23</sup> <http://www.linguateca.pt/COMPARA/index.php>

<sup>24</sup> <http://sli.uvigo.es/CLUVI/>

<sup>25</sup> [www.fflch.usp.br/dlm/comet/consulta\\_cortrad.html](http://www.fflch.usp.br/dlm/comet/consulta_cortrad.html)

## CAPÍTULO 2

### 2 METODOLOGIA

Neste capítulo, será apresentada a metodologia desta pesquisa tratando (i) do desenho do *corpus EuroParl*; (ii) do desenho do *corpus Compara*; (iii) dos procedimentos de coleta da amostra do item lexical *hearing*; (iv) do processamento dos *corpora*; e, (v) dos procedimentos para análise dos dados.

Os desenhos dos *corpora* de estudo foram feitos com base em alguns critérios propostos por Berber Sardinha (2000), Fernandes (2006) e Baker (1995), em relação ao tipo de *corpora* proposto pela autora desta dissertação, tais como a classificação do *corpus*, o número de línguas, o domínio, a direcionalidade, a restrição temporal, o modo e o tamanho do *corpus*.

#### 2.1 Desenho do Corpus EuroParl

O primeiro *corpus* a ser analisado é o *EuroParl*, que pertence à área jurídico-legislativa, contendo transcrições dos debates do Parlamento Europeu entre 1996 e 2003. É um *corpus* multilíngue, escrito em 11 línguas: alemão, dinamarquês, espanhol, finlandês, francês, grego, inglês, italiano, neerlandês, português e sueco.

Com o aumento dos países membros da União Europeia em 2004, para 25 países, eles começam a traduzir para outras línguas, que hoje, são 23 línguas oficiais traduzidas e 27 países membros. Dentre os países que fazem parte da União Europeia<sup>26</sup>, podemos destacar: Alemanha, Áustria, Bélgica, Bulgária, Chipre, Eslováquia, Eslovênia, Espanha, Estônia, Filândia, França, Grécia, Hungria, Irlanda, Itália, Letônia, Lituânia, Luxemburgo, Malta, Países Baixos, Dinamarca, Polônia, Portugal, Reino Unido, Republica Checa, Romênia, e Suécia. Alguns países, como a Macedônia, Croácia e Turquia, encontram-se em fase de negociações. Porém, será usado o par linguístico bidirecional português (PT) - inglês (EN) para investigação. O *corpus* a ser estudado possui como língua fonte tanto o inglês, como o português e/ou outra língua oficial dos países que compõem a União Europeia e fazem parte do Parlamento Europeu que foi traduzida para o inglês e/ou português.

---

<sup>26</sup> Países União Européia. Disponível em julho/2009  
<[http://pt.wikipedia.org/wiki/Uni%C3%A3o\\_Europeia](http://pt.wikipedia.org/wiki/Uni%C3%A3o_Europeia)>

O *corpus EuroParl* possui 28.521.967 palavras em inglês e 29.213.348 palavras em português. Segundo a classificação de tamanho de *corpus* proposto por Berber Sardinha (2000), o *EuroParl* é um *corpus* considerado de tamanho grande. Quanto ao modo, podemos dizer que ele é falado, pois é composto por textos transcritos dos debates do Parlamento Europeu, sendo considerado de domínio restrito-específico da área legislativa. Em relação à temporalidade, é um *corpus* sincrônico, pois se refere a um período curto de tempo que vai de 1996 a 2003.

O item lexical de busca dentro do *corpus EuroParl* é a palavra *hearing*. Foram encontradas 2.842 ocorrências na totalidade do *corpus EuroParl*. No entanto, utilizamos o *corpus EuroParl* que já se encontra alinhado na ferramenta *NATools*, que permite a busca de, no máximo, 500 ocorrências de determinado item lexical.

Como o objetivo era analisar aproximadamente 100 ocorrências aleatórias do item lexical *hearing* dentro do *corpus* legislativo para depois comparar com as ocorrências de *hearing* dentro do *corpus Compara*, foi feita uma seleção aleatória pelo *WordSmith Tools*, versão 3.0, através do qual foram obtidos 110 ocorrências desse item lexical.

A Figura 3 mostra a *Interface Natools*. Para ser feita a seleção do item lexical *hearing*, primeiramente selecionamos um dos *corpora* disponíveis na interface, que no caso foi o *EuroParl PT-EN*. Em seguida, selecionamos o número de ocorrências a serem buscadas no sistema no ícone: *Result set-size*, depois digitamos a palavra *hearing* em *Search on EN language* e por último pesquisar “*Search*”.

### Figure 3 - Interface NATools

The screenshot shows the NATools Corpora Query Interface. The search parameters are: Corpus: EuroParl-PT-EN, Search on PT language: (empty), Search on EN language: hearing, Result-set size: 500. The search results are displayed in a table with the following columns: #, %, Source Language, Target Language, and Tools.

#	%	Source Language	Target Language	Tools
1		Por que razão tivemos de passar por todo este processo para nos darmos conta de que a DG XXII estava a fazer o seu trabalho como deve ser ?	We look forward to <b>hearing</b> more of his ideas over the next few months.	[N]
2		Aquando da sua comparecência em comissão, Senhor comissário Kinnock, vimos que ainda havia 500 lugares autorizados pelo Parlamento por preencher.	In your <b>hearing</b> , Mr Kinnock, we heard that there are 500 posts, which have been authorised by Parliament, which are still vacant.	[N]
3		As propostas delineadas pelo comissário Kinnock na sua <b>audição</b> , sobretudo no domínio da formação, requerem uma resposta mais pormenorizada.	The proposals outlined by Commissioner Kinnock in his <b>hearing</b> , in particular in the field of training, need more detailed response.	[N]
4		Ficamos a aguardar com expectativa que, ao longo dos próximos meses, se pronuncie mais aprofundadamente sobre as suas ideias.	We look forward to <b>hearing</b> more of his ideas over the next few months.	[N]
5		Quanto ao comissário indigado Bolkestein, entendemos que as condições que o mesmo Presidente Prodi impôs à sua equipa o obrigam a renunciar à presidência da Internacional Liberal, por incompatibilidade, embora não pretendamos que renuncie as suas ideias, sobretudo aquelas que manifestou na sua <b>audição</b> .	With regard to Mr Bolkestein, the conditions which President Prodi himself imposed on his team, as we understand it, oblige him to renounce the Presidency of the Liberal International, because of incompatibility, and not because we want him to renounce his ideas, especially those which he demonstrated.	[N]

O *NATools*<sup>27</sup>, de responsabilidade de Lars Nygaard (The Text Laboratory) e Jörg Tiedemann (Department of Linguistics, Uppsala University), é uma ferramenta computacional para o processamento de *corpora* paralelos. Dentre outras coisas, inclui:

- Um alinhador de *corpora* ao nível da frase. Assim, cada frase do texto de partida encontra-se alinhada com o texto correspondente na tradução.

- Um extrator de dicionários probabilísticos de tradução: os dicionários probabilísticos de tradução são extraídos automaticamente a partir de *corpora* paralelos. São basicamente dicionários bilíngues, que associam a cada palavra de uma das línguas um conjunto de possíveis traduções na outra língua. Uma entrada nesse dicionário poderia ser, por exemplo, a palavra: casa (132 ocorrências); *house* (70% dessas ocorrências) e *home* (30%). Esses dicionários de tradução são chamados probabilísticos porque têm a cada tradução associada uma probabilidade, que é obtida estatisticamente a partir do *corpus* paralelo.

- Um conjunto de ferramentas para a pesquisa de *corpus* na *web*: estas ferramentas são usadas para procurar concordâncias, como no endereço: <http://linguateca.di.uminho.pt/nat>.

- Um extrator de terminologia bilíngue/paralelo: este extrator ainda se encontra em desenvolvimento, segundo Alberto Simões (2009)<sup>28</sup>. O extrator usa *corpora* paralelos, os dicionários de tradução probabilísticos e um conjunto de regras bilíngues. Pode-se fazer a busca de substantivo/adjetivo = adjetivo/substantivo, o sistema irá procurar pares de palavras que sejam substantivo adjetivo respectivamente, que tenham, de acordo com os dicionários probabilísticos, uma tradução que seja um par adjetivo/substantivo.

No decorrer da pesquisa verificou-se que existem dois sites relacionados ao Parlamento Europeu: o *site* do Parlamento Europeu onde se encontram textos traduzidos em 22 línguas, Portal do Parlamento Europeu<sup>29</sup> e 23 línguas traduzidas no Portal da União Europeia<sup>30</sup>. Considerando-se as 23 línguas existentes no portal da União Europeia, são mais de 506 combinações linguísticas possíveis, pois cada língua pode ser traduzida para outras 22 línguas.

---

<sup>27</sup> *NATools* = é uma ferramenta computacional desenvolvida por Alberto Manoel Simões atual doutorando da Universidade do Minho em Portugal.

<sup>28</sup> Alberto Simões (2009) – em um e-mail solicitando informações sobre a ferramenta *Natools*. [albie@alfarrabio.di.uminho.pt](mailto:albie@alfarrabio.di.uminho.pt) (e-mail de Simões disponível em agosto de 2009 em: [http://www.linguateca.pt/aval\\_conjunta/acetatosAvalon/avalinha.pdf](http://www.linguateca.pt/aval_conjunta/acetatosAvalon/avalinha.pdf))

<sup>29</sup> Disponível em agosto/2008 <<http://www.europarl.europa.eu/>>

<sup>30</sup> Disponível em agosto/2008 <<http://www.europa.eu/>>

Atualmente existem três sedes do Parlamento Europeu, que estão localizadas em: Bruxelas, Estrasburgo e Luxemburgo. Algumas das seções que ocorrem no Parlamento são abertas ao público.

O Parlamento Europeu também dispõe de serviços de interpretação simultânea, tradução e verificação dos textos jurídicos. O Parlamento Europeu conta com aproximadamente 430 intérpretes, que são funcionários, e mais os intérpretes externos para suprir as necessidades nessa área. A função do intérprete é “Restituir, fielmente, em todas as línguas oficiais e em tempo real...”<sup>31</sup>. A missão dos intérpretes consiste em fazer com que as reuniões decorram como se todos os participantes falassem a mesma língua, enquanto que cabe aos tradutores assegurarem as diferentes versões linguísticas dos documentos escritos.

Para a realização das atividades tradutórias, o Parlamento conta com cerca de 700 tradutores, mais os tradutores externos que são contratados quando necessário. A missão desses tradutores é traduzir para todas as línguas oficiais diversos tipos de documentos. Esses são traduzidos a partir do texto original para a língua materna do tradutor. No entanto, devido ao grande número de combinações linguísticas, nem sempre é fácil encontrar tradutores para alguns pares linguísticos, principalmente, os menos falados na Europa. Então, existem as línguas que são chamadas de *Pivot* (inglês, francês e alemão), utilizadas quando não há um tradutor para traduzir da língua fonte para a língua alvo.

Além disso, há uma grande preocupação quanto ao ensino e aprendizagem de outras línguas e ao multilinguismo: “Consagrado nos Tratados europeus, o multilinguismo reflecte a diversidade cultural e linguística da União Europeia. Graças a ele, as instituições Europeias tornam-se mais acessíveis e transparentes para os cidadãos, o que constitui uma garantia de funcionamento democrático”.<sup>32</sup>

## 2.2 Desenho do Corpus Compara

Segundo Diana Santos (2010)<sup>33</sup>, o *corpus Compara* foi desenvolvido para benefício de todos os que estudam, investigam ou trabalham com as línguas portuguesa e inglesa. Como, por exemplo falantes nati-

---

<sup>31</sup> Disponível em março de 2010 em:

<http://www.europarl.europa.eu/parliament/public/staticDisplay.do?language=PT&id=155&pageRank=4>

<sup>32</sup> Disponível em março de 2010 em:

<http://www.europarl.europa.eu/parliament/public/staticDisplay.do?language=PT&id=155>

<sup>33</sup> Disponível em março 2010 em:

<http://www.periodicos.ufsc.br/index.php/traducao/article/viewFile/5981/5685>

vos de português a aprender inglês, falantes nativos de inglês a aprender português, autores de materiais didáticos de português e inglês como língua estrangeira, tradutores, professores e estudantes de tradução (ou qualquer outro estudante iniciante ou não), lexicógrafos, engenheiros ligados ao processamento da linguagem humana para desenvolverem ferramentas computacionais, linguistas interessados no estudo da tradução e investigadores na área da tradução comparada.

O *Compara* é um *corpus* paralelo, bilíngue (inglês/português) e bidirecional, pois possui textos tanto originais em inglês e suas respectivas traduções para o português, quanto textos em português com traduções para o inglês, gratuito e de fácil acesso na *web*<sup>34</sup>. O *corpus* possui, aproximadamente, três milhões de palavras (1.358.841 palavras em língua portuguesa, incluindo originais e traduções, e 1.457.921 palavras em língua inglesa, contando com textos originais e traduções)<sup>35</sup>, o que, segundo classificação de tamanho de *corpus* proposto por Berber Sardinha (2000), é um *corpus* de tamanho médio-grande e de domínio literário.

Até o momento, o *Compara* contém 75 pares originais/traduções de textos literários publicados da África do Sul, Angola, Brasil, Estados Unidos, Moçambique, Portugal e Reino Unido, sendo que todos são traduções diretas: o português traduzido do inglês e o inglês traduzido do português. Segundo Berber Sardinha (2004), quanto à restrição temporal, o *corpus* de estudo é diacrônico, pois compreende vários períodos de tempo; já o modo apresentado é escrito, composto por textos literários.

O *Compara* é uma ferramenta que permite o estudo da tradução do inglês para o português e do português para o inglês, possibilitando, assim, não somente uma análise tradutória, mas também permite comparar as variantes do inglês ou do português de diferentes países. Além disso, o pesquisador ainda pode delimitar o *corpus*, fazendo uma busca somente por textos escritos em português brasileiro e suas traduções em inglês britânico, por exemplo.

O *Compara* apresenta tanto textos contemporâneos quanto textos antigos, o que possibilita comparar textos do mesmo original traduzidos em tempos diferentes, ou ainda delimitar qual o ano se quer incluir na pesquisa (dando margem para um estudo sincrônico ou diacrônico).

O principal objetivo do *Compara* é oferecer frases para a análise da tradução de português para o inglês e vice-versa. Não há como inves-

---

<sup>34</sup> Disponível em julho 2008 em <<http://www.linguateca.pt/COMPARA/>>

<sup>35</sup> Disponível em agosto de 2009 em: <<http://www.linguateca.pt/COMPARA/>>

tigar o capítulo inteiro, por exemplo, porque o que existe, hoje, disponível no *Compara* são apenas excertos de livros por questões de direitos autorais. Também não se encontram figuras, gráficos, diagramas ou numerações de páginas.

O alinhamento do *Compara* é feito por meio de frases. Dessa forma, cada frase do texto original encontra-se alinhada com o texto traduzido, seja uma, duas ou mais de duas frases. O *Compara* tem acesso pela *Interface do Dispara*, desenvolvida para servir de ponte entre o *IMS Corpus Workbench* e o *Compara*. Ele permite três tipos de busca: pesquisa simples, pesquisa avançada e pesquisa ultra-avançada. A busca simples pode ser feita por qualquer pessoa sem muita experiência na utilização de *corpora*. Ela permite consultar a totalidade do *Compara*, de português para o inglês e vice-versa, e os resultados são apresentados em forma de concordâncias paralelas. Basta o pesquisador digitar a palavra ou expressão em inglês ou português e acionar o botão de procura.

A pesquisa avançada serve para fazer uma pesquisa mais detalhada para que o pesquisador possa selecionar apenas partes que interessam do *Compara* para a pesquisa. Os resultados podem ser representados em forma de concordâncias paralelas ou em diversas outras formas, conforme o objetivo da pesquisa. Podem-se escolher as variantes linguísticas com as quais se pretende trabalhar, por exemplo original do português brasileiro e tradução do inglês americano, também é possível pesquisar os textos por ano de publicação. Já a busca ultra-avançada destina-se a pesquisadores que conhecem bem o *Compara* e não precisam de ajuda para pesquisar o *corpus*.

A palavra de busca no *Compára* foi *hearing* (Figura 4) com 88 ocorrências do inglês para o português encontradas em: COMPARA 13.1.9 <http://www.linguateca.pt/COMPARA/> [19-Agosto-2009]. A partir dessa seleção, foi analisado como as palavras se comportam no contexto literário.

**COMPARA - Resultados da pesquisa - Mozilla Firefox**

http://www.linguateca.pt/COMPARA/processa\_pesquisa.php

**COMPARA** This page in English

**Resultados da pesquisa**

Voltar [Imprimir](#)

Os resultados das buscas efectuadas no COMPARA podem ser usados para fins educacionais e investigação, desde que se mencione a fonte. Para citar textos específicos do corpus, selecione o código azul ao lado de cada concordância de modo a obter a sua referência completa. Para citar o COMPARA em português, use Frankenberg-Garcia, A. & Diana Santos. "COMPARA, um corpus paralelo de português e inglês na Web". *Cadernos de Tradução* IX, 2002/1. Universidade Federal de Santa Catarina, Brasil, pp 61-79. Para se referir à presente versão do corpus, escreva: **COMPARA 13.1.9** <http://www.linguateca.pt/COMPARA/> [19-Agosto-2009]

Procura: **hearing** Pedido de: **concordância em contexto**. Direcção da pesquisa: **De inglês para português**. Resultados: **88** ocorrências. Expressão de pesquisa: **"hearing"**

Descrição do corpus usado nesta procura: **1435926** palavras portuguesas, **1542762** palavras inglesas, **97723** unidades de alinhamento.

**Concordância**

<a href="#">ERDL111</a> (1775)	«Tubby, I don't believe I'm <b>hearing</b> this,» said Ollie.	-- Bolinha, não acredito no que estou a ouvir -- disse o Ollie.
<a href="#">ERDL111</a> (1883)	I don't recall <b>hearing</b> the word «recession» until a few years ago.	Só de há uns anos para cá me lembro de ouvir a palavra recessão.
<a href="#">ERDL112</a> (1775)	«Tubby, I don't believe I'm <b>hearing</b> this,» said Ollie.	-- Tubby, não acredito que estou ouvindo isso -- disse Ollie.
<a href="#">ERDL112</a> (1883)	I don't recall <b>hearing</b> the word «recession» until a few years ago.	Não consigo me lembrar de ter ouvido a palavra «recessão» até uns poucos anos atrás.

Concluído

**Figure 4 - Interface Compára**

### 2.3 Procedimentos de coleta da amostra do item lexical hearing

Nessa fase de preparação do material a ser utilizado para compor o *corpus*, serão apresentados os procedimentos de busca do item lexical a ser estudado. Os *corpora* a serem estudados não precisam ser capturados e nem alinhados, pois já se encontram prontos e disponíveis na *web*<sup>36</sup>.

A primeira etapa da seleção do material foi decidir qual item lexical capturar para o estudo. Foi decidido que seria analisado o item lexical *hearing* como amostra para o estudo proposto, pois ele aparece

<sup>36</sup>Disponível em: <<http://linguateca.di.uminho.pt/nat/nat.pl>> <<http://www.linguateca.pt/COMPARA/psimples.php>>. Acesso em agosto de 2008.

tanto em textos específicos/técnicos quanto em textos da língua em geral e é representativo dentro do contexto legislativo. O item lexical *hearing* é uma palavra polissêmica, apresenta várias traduções dependendo do contexto (Quadro 2). Por isso, a proposta é estudar as co-ocorrências, para verificar quais as padronizações que podemos encontrar ao redor do item lexical *hearing* e a influência que isso pode ter na sua tradução.

Os procedimentos da coleta da amostra dos dados incluem: (i) a obtenção de uma concordância do item lexical *hearing* no *corpus Euro-Parl* para analisar as co-ocorrências desse item lexical. Em um primeiro momento, foi feita a seleção do *corpus* na interface do *NATools* de 500 itens lexicais de *hearing*. Num segundo momento, usando a ferramenta *WordSmith Tools*, foi feita uma seleção aleatória, pedindo 1 em cada 5 palavras (das 500 palavras), onde foi obtido 110 itens lexicais; (ii) depois, uma busca de concordâncias da palavra *hearing* no *corpus Compara*, para analisar as co-ocorrências e significados dentro do contexto literário. Na interface do *Compara* foram encontradas 88 ocorrências do item lexical *hearing*. Já para o mesmo *corpus* usado na ferramenta *WordSmith Tools*, foram encontrados dois itens lexicais a mais (90 ocorrências de *hearing*).

Este estudo pode ajudar aprendizes ou tradutores a perceberem que existe uma padronização de palavras que normalmente ocorrem ao redor de um item lexical e que essa padronização pode definir o significado de um termo. Ademais, conhecer essas colocações ajuda o aprendiz a desenvolver a língua com mais fluência e naturalidade, pois os exemplos são de textos autênticos da língua a ser estudada. No Quadro 2 apresenta-se uma pequena amostra do que se pode encontrar em ambos os *corpora*.

<b>AMOSTRA DO ITEM LEXICAL “HEARING” RETIRADAS DOS CORPORA COMPARA E EUROPARL</b>	
<b>1. We look forward to <i>hearing</i> more of his ideas over the next few months .</b>	<b><i>Ficamos a aguardar com expectativa que, ao longo dos próximos meses , se pronuncie mais aprofundadamente sobre as suas ideias .</i></b>
<b>2. Even Professor Prodi declined to face an open <i>hearing</i></b>	<b><i>até o Professor Romano Prodi declinou submeter-se a uma <b>audição</b> pública .</i></b>
<b>3. With regard to the enforced <i>hearing</i> of the 79 000 appeals , the UN secretary-general cannot</b>	<b><i>No que diz respeito ao <b>tratamento</b> obrigatório de 79 000 recursos , o secretário-geral da ONU não pode</i></b>

<p><i>give precise details of the timetable or the additional staff needed to bring the hearing to a close .</i></p>	<p><i>dar uma informação exacta sobre o calendário ou a disponibilização de mais pessoal para que o procedimento de recurso possa ser levado a cabo .</i></p>
<p><i>4. The provisions require a proposal from the Commission and a hearing in the European Parliament , for the Council to take a legal decision .</i></p>	<p><i>Para que o Conselho possa tomar uma decisão legal , é necessário que haja uma proposta da comissão e que seja feita uma consulta ao Parlamento .</i></p>
<p><i>5. For this directive , the Committee on Legal Affairs and the Internal Market is proposing an advisory committee which would give the Commission the power to decide on the implementing measures after hearing the opinion of the Member States ' representatives .</i></p>	<p><i>A comissão dos Assuntos jurídicos e do Mercado Interno propõe , para esta directiva , um comité consultivo , que dá à comissão o poder de decidir as medidas de execução depois de ter recolhido o parecer dos representantes dos estados-membros .</i></p>
<p><i>6. Where the compères of most American phone-in programmes were bland, evasive, middle-of-the-road men, giving a fair hearing to all sides of the question -- endlessly patient, endlessly courteous, ultimately without convictions -- Charles Boon was violently, wilfully opinionated.</i></p>	<p><i>Enquanto os mediadores de programas similares da maioria das estações americanas eram homens comuns, evasivos, calmos e imparciais, oferecendo tempo suficiente para os dois lados se posicionarem -- com paciência e cortesia infinitas, sem qualquer convicção pessoal --, Charles Boon possuía opinião própria, formulada e defendida com violenta paixão.</i></p>
<p><i>7. This was hardly fair to Polly, who had had only two affairs in the last two years, but coming back into contact with all these good Catholics, hearing mass for the first time in ages, looking at radiant Angela and proud Dennis, and feeling sure that, incredible as it might seem, they were going to the nuptial bed as</i></p>	<p><i>Não era justo em relação à Polly, que só tinha tido dois namorados nos últimos dois anos, mas voltar a contactar com todos estes bons católicos e assistir à missa ao fim de tantos anos, ver Angela radiosa e Dennis orgulhoso e ter a certeza de que, por incrível que parecesse, eles iam para o leito nupcial genuinamente virgens fê-la sentir-se uma</i></p>

<i>genuine virgins, she herself felt distinctly Magdalenish, and began to wonder whether she should start going to mass again, perhaps even to Confession, and try to make a fresh start.</i>	<i>perfeita Madalena e começou a pensar se não deveria voltar a ir à missa, talvez até à confissão, e tentar recomeçar do zero.</i>
<i>8. Fifty-nine different people consciously misquoted T. S. Eliot's «East Coker», declaiming «What is the late December doing / With the disturbance of the spring?» in the hearing of the Americana's bell captain, to his considerable puzzlement.</i>	<i>Cinquenta e nove pessoas diferentes, conscientes de que estavam a fazer uma citação incorrecta do East Coker , de T. S. Eliot, declamaram Que anda o final de Dezembro a fazer / Perturbando a Primavera? aos ouvidos do porteiro do Americana, aumentando assim a sua já considerável perplexidade.</i>
<i>9. And what will Lizzie say to hearing that I am pregnant?</i>	<i>E que irá Lizzie dizer ao saber que estou grávida?</i>
<i>10. He told himself -- seeing their expressions, hearing snatches of their remarks in his mind, a day or two afterwards -- that he must first get his higher certificate and then he'd see.</i>	<i>Ele disse para si mesmo -- vendo as expressões deles, lembrando trechos de suas afirmativas, um ou dois dias depois -- que primeiro deveria conseguir seu diploma e daí veria o que fazer.</i>

Quadro 2: Corpora paralelos (uma amostra do Compara e Europarl)

No Quadro 2 foram selecionadas algumas frases e suas traduções (as frases de 1 a 5 foram retiradas do *NATools*, e as de 6 a 10 do *Compara*) para uma breve visualização de como a tradução de *hearing* do inglês para o português pode se comportar dependendo do contexto. Com essa pequena amostra, pode-se perceber que há uma variação considerável na tradução do item lexical *hearing* nesses dois contextos.

Após escolher o item lexical, foi feita a captura das palavras. Foi digitado o item lexical *hearing* no *Compara* e depois no *EuroParl* e depois clicou-se no ícone *Buscar*. Para fazer uma análise comparativa, precisa-se dos textos alinhados, os quais já se encontram prontos na *web*, porém o objetivo principal é analisar quais são as co-ocorrências das palavras para ver qual influência isso pode ter no significado do item lexical *hearing*. Para isso os *corpora* foram salvos no *Microsoft*

*Word for Windows* em arquivo texto (txt), e foi usado o programa de análise lexical desenvolvido por Mike Scott, *WordSmith Tools*, versão 3.0.

## 2.4 Procedimentos para Processamento do Corpus

No processamento do *corpus*, segundo Fernandes (2004), elementos computacionais devem ser utilizados para análises textuais. Para desenvolver o estudo, será utilizado o programa de análise textual *WordSmith Tools*, *software* que foi desenvolvido por Mike Scott (1996) e publicado pela *Oxford University Press*, e que se trata de um conjunto integrado de programas que permite analisar como as palavras se comportam em um texto.

O *WordSmith Tools* é composto de três ferramentas básicas para análise linguística: (i) *WordList*, (ii) *KeyWords*, e o (iii) *Concord*. O *Wordlist* elabora listas de palavras a partir de arquivos de texto, que, quando acionado, produz duas listas de palavras. Uma em ordem alfabética e outra em ordem decrescente de frequência. O *KeyWords* permite que o usuário compare uma lista de palavras de seu *corpus* de estudo (ou mais listas de vários *corpora*) com um *corpus* de referência. E o *Concord* é uma ferramenta que produz concordâncias, muito utilizado para análises linguísticas e foi utilizado para a análise do item lexical de estudo. As concordâncias são mostradas como mostra a Figura 4 abaixo. De acordo com SARDINHA (1999) são: “listagens de ocorrências de um item específico (chamado palavra de busca, palavra-chave ou nódu-lo, que pode ser formado por uma ou mais palavras), acompanhado do texto ao seu redor (co-texto)”

A preferência por esse programa é a facilidade de uso, primeiramente porque é um programa desenvolvido para o ambiente *Windows*, e, portanto, de fácil aplicabilidade, considerando que é um ambiente de utilização mundial nos dias de hoje. Uma das vantagens dessa ferramenta seria o fato de podermos observar os fenômenos da língua de forma mais rápida e segura. Outra vantagem seria calcular a frequência para obter o que é mais provável na língua. Esses cálculos de frequência podem ser feitos com o *Wordlist*. Outra ferramenta muito útil para quem utiliza *corpora* para análise da língua é o *Concord*, este instrumento produz concordâncias, que são listas de orações normalmente incompletas, nas quais uma palavra ou mais aparecem centralizadas.

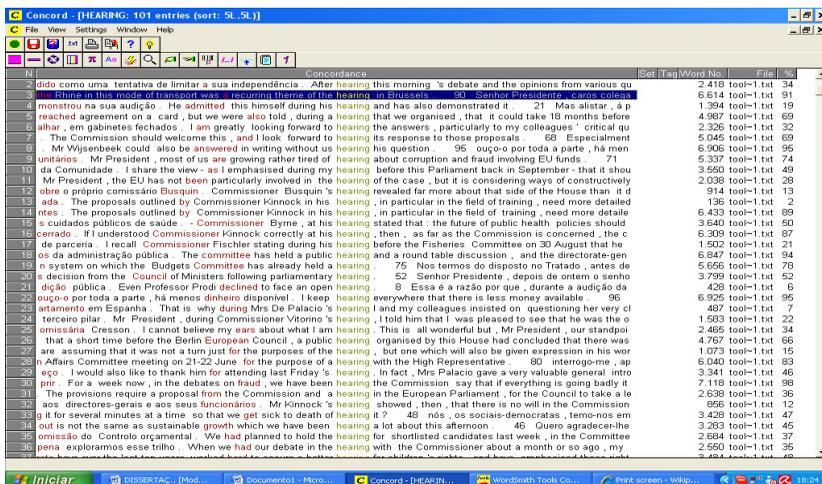


Figure 5 - Interface do WordSmith Tools

## 2.5 Procedimentos para análise de dados

A análise dos dados consiste em trabalhar o material coletado, analisando as tendências da linguagem, os padrões, as relações e as interferências. Os procedimentos para análise dos dados deste projeto de pesquisa aconteceram da seguinte forma: em um primeiro momento, foram analisadas as co-ocorrências do item lexical *hearing* e suas respectivas traduções dentro do *corpus* legislativo *EuroParl*; em seguida, foi feita uma análise do item lexical *hearing* e suas traduções dentro do contexto literário do *corpus Compara*; por último, uma análise contrastiva dos resultados finais obtidos entre os dois *corpora*: *Compara* (literário) *versus EuroParl* (jurídico/legislativo).

## CAPÍTULO 3

### 3 ANÁLISE DOS DADOS

A observação da linguagem para analisar os dados é sem dúvida muito mais eficaz do que a intuição humana. Sendo assim, a análise de *corpora* se torna imprescindível para o estudo da língua. Hunston (2002, p.39) afirma que “[...] observed language is a more valid object of study than intuition, that it matters how frequently a linguistic item occurs [...]”.<sup>37</sup> A frequência de ocorrência de um item lexical indica a importância do mesmo dentro da especialidade do corpus, por exemplo, no corpus EuroParl observa-se 2.842 ocorrências do item lexical “*hearing*”. Para estudos linguísticos considera-se que sua frequência é de grande valor investigativo, pois é uma palavra significativa dentro do contexto legislativo. O corpus é o meio mais eficaz, confiável e rápido para fazer essas análises linguísticas.

Segundo Partington (1998), o sinônimo das palavras é um membro importante para a relação da teorização existente entre as palavras. Essa organização de estabelecer sinônimos para as palavras, a fim de definir o seu significado é uma das características da maioria dos dicionários. Ao buscar no dicionário *online* Michaelis inglês/português, encontramos como correspondentes tradutórios para a palavra *hearing* as seguintes respostas: 1. *ouvido, audição, ação de ouvir*; 2. *audiência e interrogatório*. Nesse dicionário *online*, podemos encontrar dois exemplos de uso: 1. *He gave me a hearing* – ele escutou-me. 2. *Within hearing* – dentro do alcance de voz. Esses exemplos não abrangem as dificuldades que serão encontradas por um aprendiz ou por um tradutor quando se depararem com dificuldades quanto ao significado e ao padrão da língua utilizada por falantes nativos. Quando buscamos a mesma palavra no dicionário impresso Michaelis (2008) inglês/português encontramos apenas: *audição* e *audiência*, sem nenhum contexto ou frases que permitam ao tradutor uma direção quanto ao significado e uso do item lexical. O mesmo acontece com o dicionário da área jurídica de Noronha (2003), em que o único correspondente tradutório para *hearing* é *audiência*.

Por motivos de variáveis estilísticas, um aprendiz, ou tradutor experiente de uma língua estrangeira tem a necessidade de encontrar léxi-

---

<sup>37</sup> [...] a observação da língua é um objeto de estudo mais válido do que a intuição, que importa quão frequente um item linguístico ocorre [...] (tradução da autora).

cos alternativos para expressar um determinado conceito e ainda precisa saber em qual contexto um determinado sinônimo dado pelo dicionário é mais aplicável na produção de seu texto ou tradução. Partington (1998, p.29) afirma que “Comparing the collocational behaviour of so-called synonyms in concordance data may supplement dictionary information and help learners decide in what circumstances substitution of one item for another is possible in a text”<sup>38</sup>. Ou seja, além dos sinônimos, o dicionário também deve apresentar o contexto em que é frequente o uso de determinado item lexical.

### 3.1 Análise das co-ocorrências do item lexical *hearing* e suas respectivas traduções no Corpus EuroParl

A análise de co-ocorrências de um item lexical exerce notável apreço para que se possa compreender como uma palavra se comporta dentro de um determinado contexto, quais padrões se podem esperar ou quais palavras aparecem ao redor do item lexical.

Para obter essas *colocações* ou *coligações*, usa-se um concordanciador, *WordSmith Tools*, um programa que pesquisa no *corpus* um determinado item lexical que o pesquisador pretende analisar. Para isso, o pesquisador digita a palavra no campo de busca, o resultado dessa concordância é uma lista com frases (cotexto), a palavra que aparece centralizada na tela do computador é chamada de nóculo de busca.

Ao observar as palavras que acompanham o item lexical *hearing* por meio da ferramenta *Concord* nos textos de debates do Parlamento Europeu foi possível identificar todos os co-ocorrentes do nóculo *hearing*.

Faz-se necessário salientar que foram analisados 110 ocorrências do item lexical *hearing* no *corpus EuroParl*, assim como 90 ocorrências do item lexical *hearing* no *corpus Compara*. No entanto, a autora decidiu por descrever apenas uma pequena amostra dessa análise para a leitura não ficar cansativa para o leitor. A Tabela 2 apresenta o número de cada co-ocorrente presente no *corpus EuroParl* do item lexical *hearing* e a porcentagem dentro do *corpus*.

Tabela 2: Lista de Co-ocorrências do item lexical *hearing* (Corpus EuroParl).

---

<sup>38</sup> Comparando o comportamento colocacional dos chamados sinônimos em concordâncias podem complementar dados do dicionário e ajudar aprendizes a decidir em qual circunstância a substituição de um item lexical por outro é possível em um texto. (Tradução da autora).

	<b>CO-OCORRÊNCIAS do item lexical “hearing”</b>	<b>NÚMEROS DE CO- OCORRÊNCIAS</b>	<b>TOTAL %</b>
1.	during the hearing	10	9,09%
2.	after hearing	9	8,18%
3.	held a hearing	5	4,55%
4.	public hearing	3	2,73%
5.	hearing about	4	3,64%
6.	look foward to hearing	5	4,55%
7.	on hearing/ hearing on	6	5,45%
8.	(verb) to be + Hering	16	14,54%
9.	hearing organized by	3	2,72%
10.	of a/the hearing	11	10%
11.	at his/our/my/a hearing	12	10,9%
12.	the hearing	5	4,54%
13.	from the hearing/ hear- ing from	6	5,45%
14.	to the hearing	3	2,72%
15.	Outras ocorrências menos frequentes	12	10,90
	<b>TOTAL</b>	<b>110</b>	<b>100%</b>

Observa-se que o item lexical *hearing* acompanhado de outros itens lexicais dentro do contexto jurídico/legislativo é frequentemente traduzido por *audiência* ou *audição*, como nos exemplos: *after the hearing* traduzido por *depois da audiência*, *during a hearing* traduzido por *numa audiência*, *held a hearing* traduzido por *realizamos uma audiência*, *public hearing* em *audição pública*, *a hearing on* por *audição sobre*, *from the excellent hearing held by* traduziu-se por *da excelente audiência realizada pela*, *the hearing organized by* traduzido por *a audiência organizada na*, etc. Apesar dessa constatação temos outros exemplos em que o item lexical *hearing* apresenta outros correspondentes tradutórios. Para melhor visualização, seguem alguns exemplos analisados dentro do contexto jurídico/legislativo.

Ao observar as palavras que acompanham o item lexical *hearing*, nos textos de debates do Parlamento Europeu encontrou-se 10 co-ocorrências de *during + hearing*, o que corresponde a 9,09% do total das ocorrências. Esse número é significativo dentro do contexto dos textos transcritos do Parlamento Europeu. As ocorrências e co-ocorrências do item lexical *hearing* são apresentadas no Quadro 3.

LINHAS DE CONCORDÂNCIA			
1	Already, during his	hearing	last month, he had stated that Turkey's candidacy...
2	And yet, during the	hearing	of 29 October 1997, Commissioner Monti recognized that
3	During that the	hearing	on human rights was agreed, which we obviously...
4	I recall Commissioner Fischler stating during his	hearing	before the Fisheries Commission is concerned, the ....
5	Mr. Barnier even complacently stressed, during his	hearing	last week, that the process of preparing the Charter...
6	Mr. President, during Commissioner Victorino's	hearing	, I told him that I was pleased to of the agenda.
7	During the public	hearing	last November, when culture and cohesion policy...
8	The President of the Court of Auditors, during a	hearing	of the Subcommittee on Monetary Affairs four...
9	Presented himself as a euro-critic and euro-sceptic, and during the	hearing	he made a 180° turn.
10	Especially those which he demonstrated during his	hearing	

Quadro 3: Ocorrências e co-ocorrências do item lexical *during* + *hearing* (Corpus EuroParl).

Observa-se que a palavra *hearing* acima citada, aparece acompanhada de: *during* + pronome (*his*), *during* + artigo (*a/the*) e *during* + substantivo (*Commissioner Victorino/ the public*). Outra ocorrência frequente é o uso de tempo ao redor do item lexical, assim como: *during his hearing last month*; *during his hearing before the Fisheries Committee on 30 August*; *during his hearing last week*; *During the public hearing last November*; *during a hearing of the Subcommittee on Monetary Affairs four months ago*.

Percebe-se também que todas as ocorrências de *hearing* analisadas acima, têm a função de substantivo. Nesse contexto, o item lexical *hearing* foi traduzido por *audição*. Como apresentado no Quadro 4.

<i>Lingua Portuguesa</i> ( <i>Texto A</i> )	<i>Lingua Inglesa</i> ( <i>Texto B</i> )
(1A) <u>já na sua <b>audição</b> , no mês passado</u> , o senhor comissário tinha declarado que a candidatura da Turquia deveria ser sujeita a nova apreciação em Dezembro próximo, no Conselho de Helsínquia , e que, pessoalmente , a apoiaria .	(1B) <u>Already , during his <b>hearing</b> last month</u> , he had stated that Turkey 's candidacy must be re-examined next December , at the Helsinki Council , and that he , personally , would support it .
(2A) lembro-me do senhor comissário Fischler afirmar <u>durante a sua <b>audição</b></u> , perante a comissão das Pescas a 30 de Agosto , que desejava estabelecer uma estreita cooperação com o Parlamento e as comissões permanentes do Parlamento num trabalho de parceria .	(2B) I recall Commissioner Fischler stating <u>during his <b>hearing</b></u> before the Fisheries Committee on 30 August that he wished to cooperate closely with Parliament and the standing committees of Parliament in a working partnership.
(3A) O senhor comissário Barnier, aliás de forma complacente, frisou <u>a- quando da sua <b>audição</b></u> na semana passada, que o processo de preparação da Carta deveria ligar-se ao processo de constitucionalização dos tratados .	(3B) Mr Barnier even complacently stressed, during his <b>hearing</b> last week, that the process of preparing the Charter should be brought together with the process of constitutionalising the Treaties.

Quadro 4: Correspondentes Tradutórios do item lexical *during* + *hearing* (Corpus EuroParl).

O fato da ocorrência de *during* + *hearing* ser frequente dentro do *corpus* do Parlamento Europeu pode ser devido ao fato de ser um texto que representa a fala dos deputados, que costumam narrar fatos ocorridos durante um determinado período de tempo. Como no exemplo “Ora, como deu a entender o senhor comissário Frits Bolkestein, **durante a audição do passado dia 6 de Setembro..**” traduzido por “It may be supposed, for example, as Commissioner Bolkestein gave us to understand during his **hearing** on 6 September last year” (Corpus EuroParl).

Outra ocorrência frequente, 9 ocorrências (8,18% das 110 ocorrências), nos textos do Parlamento é o uso de *after + hearing*, conforme demonstrado no Quadro 5.

LINHAS DE CONCORDÂNCIA			
11	Of the Treaty on European Union and, at all events, after	hearing	Parliament's opinion.
12	And after	hearing	what Mr. Fischler had to say I can only endorse....
13	I do think that the Bureau, in its wisdom, after	hearing	this confession this morning would wish to ...
14	... for a resolution which was not drafted in haste, straight after	hearing	the outcome of the conference, but ...
15	After	hearing	from us, the Council long...
16	... to take on our foreign policy after	hearing	that statement about one of the countries...
17	Basing myself on the preparatory work, and after	hearing	the two rapporteurs, Stanislay Tillich and John ....
18	Oostlander report waited until after that	hearing	or, indeed, that the hearing be brought forward...
19	So, especially after the	hearing	we have held, and after the strong warning...

Quadro 5: Ocorrências e co-ocorrências do item lexical *after + hearing* (Corpus EuroParl).

Nos exemplos citados acima, o item lexical *hearing* co-ocorre, de forma frequente, com a palavra *after*. Entende-se que o uso frequente de *after* corresponde a um texto que expressa ações que aconteceram e/ou ainda estão por acontecer, ou seja, discussões de assuntos a serem resolvidos ou que já foram resolvidos no Parlamento Europeu.

Os correspondentes tradutórios do item lexical *hearing* que precedem ao item lexical *after* não seguem as mesmas regras que *during +*

*hearing*, em que *hearing* apresentava função de substantivo. No exemplo (11) do Quadro 5, o tradutor optou pelo correspondente tradutório *parecer*; nos exemplos (12), (13) e (16), o correspondente tradutório de *hearing* foi o verbo *ouvir*; no (14), *ter conhecimento*; no (15), *informações*; no exemplo (17), *ter ouvido*; no (18), *audiência*; e no exemplo (19), *audição*. Ao observar os textos alinhados no Quadro 6 (português/inglês), tem-se a impressão que o processo de tradução é simples. No entanto, o aprendiz ao deparar com o termo *hearing* pode apresentar grandes dificuldades tradutórias devido às variantes tradutórias e a complexidade do item lexical quando contextualizado. Os exemplos demonstrados dentro de um determinado contexto são ferramentas de grande valor para aprendizes ou até mesmo para tradutores experientes, pois permitem uma maior visualização do significado real dos termos. A seguir, são apresentados alguns exemplos do *corpus* paralelo (inglês-português) *EuroParl* no Quadro 6.

<i>Língua Portuguesa</i> ( <i>Texto A</i> )	<i>Língua Inglesa</i> ( <i>Texto B</i> )
(4A) Ora , a minha proposta é muito simples , retoma o texto da comissão e acrescenta que o Conselho pode estabelecer medidas adequadas deliberando por maioria qualificada , por proposta da comissão ou com base numa recomendação do Parlamento Europeu , adoptada nos termos do artigo j.7 do Tratado da união Europeia e , em qualquer caso , <u>após o parecer do Parlamento Europeu .</u>	(4B) My proposal is very simple : the Commission 's text should be adopted with the addition that the Council may establish appropriate measures , deciding by a qualified majority on the Commission 's proposal or on the basis of a recommendation made by the European Parliament , adopted in accordance with Article j.7 of the Treaty on European Union and , at all events , <u>after <b>hearing</b> Parliament 's opinion .</u>
(5A) De resto , face ao que me foi dado <b>ouvir</b> do senhor comissário Fischler , apenas poderei apoiar a posição da comissão , nomeadamente de observar os Acordos de florença .	(5B) And after <b>hearing</b> what Mr Fischler had to say I can only endorse the Commission 's position that we must stick to what was agreed in Florence .
(6A)	(6B)

<p>Senhor Presidente, gostaria de fundamentar o pedido de adiamento da declaração da comissão referente à conferência de Buenos Aires sobre o clima, invocando o facto de os colegas que participaram nesta conferência em representação do Parlamento e eu próprio termos debatido o assunto, e considerarmos ser mais prudente preparar cuidadosamente tal discussão nas respectivas comissões. então, poder-se-á terminar com uma resolução que será elaborada de forma exacta e cuidadosa, e não à pressa, <u>logo após se ter conhecimento dos resultados.</u></p>	<p>Mr. President , let me justify my motion to postpone the Commission statement on the Buenos Aires climate conference on the grounds that I and the Members who took part in this conference on behalf of Parliament decided , after talking the matter over , that it would be more sensible to prepare carefully for this discussion in the appropriate committees ; that would allow us to conclude with a motion for a resolution which was not drafted in haste , <u>straight after hearing the outcome of the conference</u> , but was based on careful and detailed preparation.</p>
---	---

Quadro 6: Correspondentes Tradutórios do item lexical *after + hearing* (Corpus EuroParl).

O item lexical *held* seguido de *hearing* ou *hearing* seguido de *held* é outra co-ocorrência frequente no *corpus* do Parlamento Europeu. Foram encontradas 5 ocorrências que equivalem a 4,55% das 110 ocorrências, conforme demonstradas no Quadro 7.

<i>LINHAS DE CONCORDÂNCIA</i>			
20	I would recall that the Committee on Agriculture and Rural Development held a	<b>hearing</b>	on mountain regions, and in particular...
21	Budgets Committee has already held a	<b>hearing</b>	.
22	Finally, we held a	<b>hearing</b>	in March in the European Parliament with a senior...
23	...the subcommittee on Human Rights	<b>hearing</b>	in Brussels on the question of impunity...

		held a	
24	... the various sectors which were consulted by the Commission at a	Hearing	held in September 1995...

Quadro 7: Ocorrências e co-ocorrências do item lexical *held* + *hearing* (Corpus EuroParl).

Os exemplos mostram que *held a hearing* são, em todos os casos traduzidos, por: *realizar uma audiência*, *efetuar uma audiência* ou *ter lugar a uma audiência* (no sentido de acontecer algo). Também se percebe o uso de lugar ou data como co-ocorrência de *held a hearing*, como nos exemplos: *held a hearing on mountain regions*, *held a hearing in March*, *held a hearing in Brussels*, etc. No Quadro 8, apresentam-se alguns exemplos da tradução de *held a hearing*.

<i>Língua Portuguesa</i> ( <i>Texto A</i> )	<i>Língua Inglesa</i> ( <i>Texto B</i> )
(7A) Lembro que, sobre os problemas da montanha e, em especial, da agricultura de montanha, <u>realizámos uma audiência</u> em 27 de Janeiro de 1997, no âmbito dos trabalhos da comissão da Agricultura e do Desenvolvimento Rural.	(7B) I would recall that the Committee on Agriculture and Rural Development <u>held a hearing</u> on mountain regions, and in particular on hill farming, on 27 January 1997.
(8A) Neste contexto, há que colocar energeticamente em prática o sistema de inscrição no orçamento, sistema relativamente ao qual a comissão dos orçamentos já <u>efectuou uma audiência</u> ao nível desta câmara e dos grupos políticos.	(8B) So it is now a question of actively implementing - both in the House and within the groups - the budgetisation system on which the Budgets Committee has already <u>held a hearing</u> .
(9A) Senhor Presidente, em 30 e 31 de Outubro <u>teve lugar em Bruxelas uma audiência</u> organizada pela sub-comissão « Direitos do Homem »,	(9B) Mr President, on 30 and 31 October the Subcommittee on Human Rights <u>held a hearing in Brussels</u> on the question of impunity and

sobre a questão da impunidade - a necessidade de uma resposta internacional -, um tema deveras difícil.	the need for an international answer - a really difficult question.
---	---

Quadro 8: Correspondentes Tradutórios do item lexical *held + hearing* (Corpus EuroParl).

No Quadro 9, há três exemplos de co-ocorrências do item lexical *hearing + public*, que corresponde a 2,73% das 110 ocorrências do *corpus EuroParl*.

LINHAS DE CONCORDÂNCIA			
25	people think our resolution of September is enough, that the public	hearing	with the Algerian authorities and civil society.
26	...during the public	hearing	last November, when culture and cohesion policy...
27	... to organizing the important public	hearing	which has allowed us to see things more...

Quadro 9: Ocorrências e co-ocorrências do item lexical *public + hearing* (Corpus EuroParl).

Ao pesquisar o item lexical *hearing* no dicionário monolíngue da língua inglesa *Longman Advanced American Dictionary (2005)*, há uma explicação para o uso de *public + hearing* isso é, “*a meeting of a court or special committee to find out the facts about a case: a public hearing*”. Contudo, não há exemplos com o colocado *public*. O exemplo apresentado é *a fair*, como no exemplo: “*give somebody a fair hearing to give someone an opportunity to explain their actions, ideas, or opinions*”. A autora usou este exemplo para ilustrar que os dicionários nem sempre trazem o contexto esperado abrindo espaço para uma investigação mais precisa em um corpus especializado.

O correspondente tradutório apresentado no *corpus EuroParl* para *public hearing* é *audiência/ audiência pública*. Como apresentado no Quadro 10.

<i>Língua Portuguesa</i> ( <i>Texto A</i> )	<i>Língua Inglesa</i> ( <i>Texto B</i> )
(10A) há quem pense que a nossa resolução do mês de Setembro é suficiente, <u>que a <b>audição pública</b></u> com as autoridades argelinas e a sociedade civil, prevista para 25 de Novembro, poderia ficar comprometida se nos exprimirmos hoje, que a delegação parlamentar ad hoc poderia ser anulada.	(10B) Some people think our resolution of September is enough, <u>that the <b>public hearing</b></u> with the Algerian authorities and civil society, scheduled for 25 November, could be compromised if we express ourselves today, that the ad hoc parliamentary delegation could be cancelled.
(11A) No decorrer <u>da <b>audição pública</b></u> de Novembro passado, em que se falou de cultura e de política de coesão, surgiu a interessante sugestão de um objectivo 7 dedicado à cultura, de um objectivo de tipo horizontal.	(11B) During <u>the <b>public hearing</b></u> last November, when culture and cohesion policy were discussed, the interesting suggestion emerged of an objective 7 devoted to culture, a horizontal objective.
(12A) Nem sempre estivemos de acordo, cara colega, mas investiu bastante energia a fim de fazer avançar este dossier, de organizar a importante <u><b>audição pública</b></u> que Nos permitiu sermos esclarecidos, De Preparar As deliberações da nossa comissão que se reuniu doze vezes para discutir este Assunto , para procurar entre nós os termos de um acordo .	(12B) We have not always been in agreement, Mrs mosiek-urbahn, but you have devoted a great deal of energy to furthering this case, to organizing the important <u><b>public hearing</b></u> which has allowed us to see things more clearly, to prepare the deliberations of our committee which has met twelve times on this subject to seek amongst ourselves the terms of an agreement.

Quadro 10: Correspondentes Tradutórios do item lexical *public + hearing* (Corpus EuroParl).

Outra co-ocorrência que aparece no *corpus é hearing about*, com 4 ocorrências, o que corresponde a 3,64 % das 110 ocorrências, como se observa no Quadro 11.

LINHAS DE CONCORDÂNCIA			
28	Commissi- sioner, although you are already	hearing	about the red-green ma- jority, I do hope that you...
29	In other contexts we are always	hearing	about the latest progress in genetic engineering.
30	We are always	hearing	about assistance to smes but no Concil Presiden- cy...
31	I would not wish to deprive you of	hearing	about the feelings I experi- enced over the weekend.

Quadro 11: Ocorrências e co-ocorrências do item lexical *hearing* + *about* (Corpus EuroParl).

No exemplo (28) do Quadro 11, *hearing about* é traduzido como *oiça falar da* (ouça falar da), ou seja, *mesmo que já tenha ouvido falar...* (although you are already hearing about...). Já no exemplo (29), o tradutor usou *salientados* e depois reforçou o mesmo sentido com *chamar a atenção*, como demonstra o exemplo: *salientados, aquilo que se procura chamar a atenção*. No exemplo (30), o correspondente tradutório é *fala constantemente*, e no exemplo (31), o tradutor optou por *expressir*. Como mostra alguns exemplos apresentado no Quadro 12.

<i>Língua Portuguesa</i> ( <i>Texto A</i> )	<i>Língua Inglesa</i> ( <i>Texto B</i> )
(13A) Senhora comissária , mesmo que já <u>oiça falar da</u> maioria verde-vermelha , espero que aceite propostas sensatas da outra parte e , como sou sindicalista , espero que nos possamos entender pelo menos a esse nível .	(13B) Commissioner , although you are already <u>hearing about</u> the red-green majority , I do hope that you accept sensible proposals from the other side . As I am a trade unionist , I hope that this is at least one level on which we can agree .
(14A) Os aspectos que são <u>salientados</u> , <u>aquilo para que se procura chamar a atenção</u> é para os mais recentes	(14B) In other contexts <u>we are always hearing about</u> the latest progress in genetic engineering .

progressos da manipulação genética .	
(15A) Senhor Presidente , para finalizar este trio neerlandês , não quero abster-me de lhe <u>exprimir</u> os meus sentimentos este fim-de-semana .	(15B) Madam President , as I round off this Dutch trio , I would not wish to deprive you <u>of hearing</u> about the feelings I experienced over the weekend . I must say it was difficult to suppress a certain amount of despondency .

Quadro 12: Correspondentes Tradutórios do item lexical *hearing* + *about* (Corpus EuroParl).

Foram encontradas 5 ocorrências da expressão *look forward to hearing* no contexto jurídico/legislativo, que representam 4,55% das 110 ocorrências apresentadas no Quadro 13.

LINHAS DE CONCORDÂNCIA			
32	And I am also looking forward to	<b>hearing</b>	them speak now.
33	I look forward to	<b>hearing</b>	the Commission's proposal regarding the ...
34	I look forward to	<b>hearing</b>	some appropriate proposals from the Commission...
35	I hope forward to	<b>hearing</b>	colleagues contributions and to hearing Commissioner Pappou...
36	We look forward to	<b>hearing</b>	more of his ideas over the next few months.

Quadro 13: Ocorrências e co-ocorrências do item lexical *look forward to hearing* (Corpus EuroParl).

As traduções encontradas para *look forward to hearing* foram: *espero ouvi-los, gostaria de ouvir, aguardo com interesse, aguardo com expectativa, ficamos a aguardar com expectativas*. Nos dicionários da

*Oxford (2000) e Longman (2005)*, que apresentam exemplos baseados em *corpus*, foram encontrados correspondentes tradutórios para *look forward to*, como nos exemplos: *I'm looking forward to the weekend* ou *We are really looking forward to seeing you again*, mas não houve exemplo utilizando o item lexical *hearing*. Apresentam-se alguns exemplos de correspondentes tradutórios de *look forward to hearing* dentro de seus contextos no Quadro 14.

<i>Língua Portuguesa</i> ( <i>Texto A</i> )	<i>Língua Inglesa</i> ( <i>Texto B</i> )
(16A) <u>E espero ouvi-los</u> agora.	(16B) And I am also looking forward to <b>hearing</b> them speak now.
(17A) <u>Aguardo com interesse propostas</u> da comissão que vão neste sentido e espero que consigamos obter rapidamente medidas concretas neste importante sector.	(17B) I look forward to <b>hearing</b> some <u>appropriate proposals</u> from the Commission and hope that in this important area it will prove possible to achieve some concrete measures quickly.
(18A) <u>Ficamos a aguardar com expectativa</u> que , ao longo dos próximos meses , se <u>pronuncie</u> mais aprofundadamente sobre as suas ideias.	(18B) <u>We look forward to</u> <b>hearing</b> more of his ideas over the next few months .

Quadro 14: Correspondentes Tradutórios do item lexical *look forward to hearing* (Corpus EuroParl).

Nos exemplos do Quadro 15, temos as *coligações* de *hearing on* ou *on hearing*, essas *coligações* são bem frequentes no *corpus EuroParl*, representando 5,45% das 110 ocorrências analisadas. Segundo Tagnin (2005, p. 31), o uso de preposições é uma das grandes dificuldades dos aprendizes de qualquer língua, principalmente porque, às vezes, não há uma explicação plausível para o fato de ocorrer uma determinada preposição em vez de outra.

<i>LINHAS DE CONCORDÂNCIA</i>			
37	Secondly, I note, on	Hearing	what she has just said, that the Commission ...
38	Moreover, we keep on	Hearing	that there is not enough money for development aid.
39	...that the European Parliament should invite these experts to a	Hearing	on safety standards at the plant.
40	It is not a custody	Hearing	on children.
41	In 1995 Mr. Collins chaired Parliament's	Hearing	on water policy that gave birth...

Quadro 15: Ocorrências e co-ocorrências do item lexical *on hearing* e *hearing on* (Corpus EuroParl).

No exemplo (37) do Quadro 15, *on hearing* é traduzido por *ao ouvir*. Já no exemplo (38), *we keep on hearing*, o termo *keep on* é um *phrasal verb* que significa *continuar fazendo* o correspondente traduzido utilizado pelo tradutor foi *ouve-se constantemente dizer*. Nos exemplos (39), (40) e (41), a preposição *on* ocorre após o item lexical *hearing* e são traduzidos por *uma audição sobre* e *audição para*, como são apresentados no Quadro 16.

<i>Lingua Portuguesa</i> ( <i>Texto A</i> )	<i>Lingua Inglesa</i> ( <i>Texto B</i> )
(19A) Em segundo lugar, <u>constato, ao ouvir</u> o que acaba de nos ler, que a comissão rejeita sistematicamente as alterações politicamente significativas que foram apoiadas hoje.	(19B) Secondly, <u>I note, on hearing what she has just said</u> , that the Commission systematically rejects all the politically significant amendments which have been supported today.
(20A) Além disso, <u>ouve-se constantemente dizer</u> que não há dinheiro suficiente para a ajuda ao desenvolvi-	(20B) Moreover, <u>we keep on hearing</u> that there is not enough money for development aid.

mento .	
(21A) Proponho que o Parlamento Europeu <u>convide estes peritos para uma <b>audição</b> sobre normas de segurança.</u>	(21B) I propose that the European Parliament should <u>invite these experts to a <b>hearing on</b> safety standards at the plant.</u>

Quadro 16: Correspondentes Tradutórios do item lexical *on hearing e hearing on* (Corpus EuroParl).

No Quadro 17, temos ocorrências de *from + hearing* ou *hearing + from*. De acordo com a classificação de Tagnin (2005, p.31), a ocorrência de *from + hearing* é chamada de *coligação*. Pode ser uma *coligação* de uma *preposição* mais um *substantivo*, como no exemplo: *Apart from the hearing in the Subcommittee on Monetary Affairs [...]*, traduzida por *Porque, além da audiência realizada na subcomissão dos assuntos monetários [...]*; ou um *verbo* mais uma *preposição*, como em *That is why you have been hearing from Belgian, Italian, French and Dutch colleagues [...]* traduzida por *É por isso que têm estado a ouvir intervenções de colegas belgas, italianos, franceses e holandeses [...]*. No Quadro 17, são apresentadas essas concordâncias.

LINHAS DE CONCORDÂNCIA			
42	... and from what we are	hearing	, not just the MED programmes, either.
43	... as many Members will know from the excellent	hearing	held by the Committee on Transport and Tourism...
44	Apart from the	hearing	in the Subcommittee on Monetary Affairs.
45	After	hearing	from us, the Council long since amended...
46	That is why you have been	hearing	from Belgian, Italian, French and Dutch colleagues...
47	I'm also worried because I have also been	hearing	this view from a number of central bankers this ...

Quadro 17: Ocorrências e co-ocorrências do item lexical *from hearing* e *hearing from* (Corpus EuroParl).

No Quadro 18, apresentam-se algumas soluções de traduções que o aprendiz, ou o tradutor, poderia encontrar para a tradução de *from hearing* ou vice-versa em um contexto jurídico/legislativo.

<i>Lingua Portuguesa</i> ( <i>Texto A</i> )	<i>Lingua Inglesa</i> ( <i>Texto B</i> )
(22A) <u>Mas segundo informações que nos chegaram</u> , o Conselho modificou há muito tempo a base textual , pelo que possivelmente estamos a exigir algo que deixou de ser objecto da discussão .	(22B) <u>After hearing from us</u> , the Council long since amended basic text , so that we may be calling for something which is no longer up for discussion at all .
(23A) A maior parte dessas acções de primeira prioridade já foram concluídas, como é do conhecimento de muitos dos membros do Parlamento depois <u>da excelente audição realizada pela</u> comissão dos Transportes e do Turismo no início deste ano .	(23B) Most of the first-priority actions have already been completed , as many Members will know <u>from the excellent hearing held by</u> the Committee on Transport and Tourism at the beginning of this year .
(24A) Porque, <u>além da audição realizada na subcomissão dos Assuntos monetários</u> , foram centenas as cartas e os e-mails que enviei às mais diversas entidades ligadas ao sector do turismo em todo o mundo .	(24B) <u>Apart from the hearing in the Subcommittee on Monetary Affairs</u> , I sent hundreds of letters and e-mails to the widest possible range of bodies connected with the tourism sector throughout the world .

Quadro 18: Correspondentes Tradutórios do item lexical *from hearing* e *hearing from* (Corpus EuroParl).

Após a análise de todos esses dados, foi possível verificar de forma quantitativa e qualitativa quais correspondentes tradutórios foram encontrados no *corpus EuroParl*. O item lexical de maior frequência foi

o substantivo *audição* e *audiência* com 55,45% das ocorrências, depois o verbo *ouvir* com 19,09% das ocorrências, os itens lexicais *assistir*, *consulta*, *exprimir*, *surdos*, *atenção*, *informação* e *falar* ocorreram duas vezes cada um (1,81%), e 2,7% dos casos foram *omissos*. As demais ocorrências obtiveram apenas uma ocorrência cada um, porém não significa que elas sejam menos importantes, apenas caracteriza que dentro do *corpus Europarl* eles não tiveram grande número de ocorrências, por exemplo: *audição* no sentido de aparelho auditivo e o verbo *escutar* tiveram apenas uma ocorrência dentro *corpus Europarl*. Na Tabela 3, apresentam-se os correspondentes tradutórios encontrados no *corpus*.

**Tabela 3: Correspondentes Tradutórios (total) do item lexical *hearing* (Corpus EuroParl).**

<i>CORRESPONDENTES TRADUTÓRIOS</i>	<i>NÚMERO DE OCORRÊNCIAS</i>	<i>%</i>
Ouvir	21	19,09
Audição	52	47,27
Audiência	9	8,18
Assistir	2	1,81
Receber	1	0,90
Registro	1	0,90
Consulta	2	1,81
Exprimir	2	1,81
Surdos	2	1,81
Audição (ouvido)	1	0,90
Encontro	1	0,90
Atenção	2	1,81
Escutar	1	0,90
Ter conhecimento	1	0,90
Informação	2	1,81
Falar (oiça falar)	2	1,81
<b>Casos omissos</b>	3	2,72

### **3.2 Análise das co-ocorrências do item lexical *hearing* e suas respectivas traduções no Corpus Compara**

Ao observar os itens lexicais que co-ocorrem com *hearing* podemos destacar alguns que ocorrem com maior frequência dentro do *cor-*

*pus Compara*, assim como: *on hearing this evidence, hard of hearing, she was/he was/you were hearing, hearing the, hearing from, hearing of, hearing about, his/her/their hearing, difficulty in hearing all, upon hearing, hearing something/someone, in the hearing of, after hearing, in the hearing, he kept hearing, seeing and hearing, to hearing*. Conhecer quais palavras co-ocorrem com o item lexical ajuda a escrever melhor e usar frases de forma que soe mais natural para um nativo.

Analisar as co-ocorrências de um item lexical implica em um estudo da tradução desse item lexical dentro do contexto, pois o significado de cada palavra está de certa forma, ligado ao meio em que a palavra se encontra. Para melhor demonstrar isso, foram analisados os *colocados* e *coligações* dos itens lexicais e posteriormente algumas possíveis traduções que vão servir de amostra do *corpus* analisado. Na Tabela 4, são apresentados os co-ocorrentes do item lexical *hearing* dentro do *corpus Compara*.

**Tabela 4: Co-ocorrências do item lexical *hearing* (Corpus Compara).**

<i>CO-OCORRENTES do item lexical “hearing”</i>	<i>NÚMERO DE CO- OCORRÊNTES</i>	<i>%</i>
(verb to be) + hearing	10	11,11%
on hearing	8	8,88%
hard of hearing	2	2,22%
in hearing	5	5,55%
upon hearing	2	2,22%
(pronoun- his/their/our/your) hearing	12	13,33%
to hearing	3	3,33%
just hearing	2	2,22%
of hearing	7	7,77%
from hearing	2	2,22%
outros	37	41,11%
<b>TOTAL</b>	<b>100</b>	<b>100%</b>

Na Tabela 5 pode-se constatar que aproximadamente 70% dos itens lexicais de *hearing* foram traduzidos por um *verbo*, e que estes verbos variam de acordo com os colocados.

**Tabela 5: Correspondentes Tradutórios do *Corpus Compara***

<b>CORRESPONDENTES TRADUTÓRIOS</b>	<b>NÚMERO DE OCORRÊNCIAS</b>	<b>%</b>
Ouvir	44	48,88
Omissos	4	4,44
Escutar	8	8,88
audição (aparelho auditivo)	2	2,22
ouvido (verbo ouvir)	8	8,88
Falar	2	2,22
vivendo (verbo viver)	1	1,11
ter conhecimento	1	1,11
Pressentir	1	1,11
Presença	1	1,11
Saber	1	1,11
Assistir	1	1,11
Lembrando	1	1,11
Adivinhando	1	1,11
em sessão	1	1,11
Tímpano	1	1,11
por perto (within hearing)	1	1,11
uma justa oportunidade/ oferecendo tempo suficiente (a fair hearing)	2	2,22

Os primeiros exemplos analisados foram as coligações do verbo *to be* + *hearing*, como aparecem no Quadro 19.

<b>LINHAS DE CONCORDÂNCIA</b>			
1	Tubby, I don't believe I'm	<b>hearing</b>	this,» said Ollie.
2	...it was the French horns we should have been	<b>hearing</b>	now, not the clarinets...
3	I've been	<b>hearing</b>	about your latest.
4	I've been	<b>hearing</b>	all about it.
5	... as if you were	<b>hearing</b>	an orchestra playing its instruments.
6	...it is almost as if we	<b>hearing</b>	Your Royal Highness

	were		should know
8	Helen looked into my eyes, frightened, as if she were	hearing	some sinister bird's cry at her doorstep.
9	Like the chief, like July, like everyone, she was	hearing	him say what he and she had always said, ...
10	...but I couldn't believe what I was	hearing	from Petrus.
11	I was	hearing	the most unexpected farewell in my life

Quadro 19: Co-ocorrências do item lexical *hearing* + *verb to be* (Corpus Compara).

No exemplo (1) do Quadro 20, a frase *I don't believe I'm hearing this* o item lexical de estudo *hearing* foi traduzido pelo verbo *ouvir* no infinitivo. Houve mudanças na estrutura da frase, a primeira mudança foi a transformação do verbo que estava no gerúndio para o infinitivo, também há o acréscimo no texto alvo de *no que* e a perda do item lexical *this* do texto fonte.

No exemplo (2) *it was the French horns we should have been hearing now* o tradutor optou por traduzir *hearing* por *ouvir-se*, percebe-se que *French* foi omitido do texto original e *now* foi traduzido por *naquele momento*.

No exemplo (3) *I've been hearing about your latest* o item lexical *hearing* foi traduzido por *ouvido*. O tradutor acrescentou no texto de chegada a palavra *falar*, a fim de deixar o texto mais natural aos leitores. Outro detalhe é que a palavra *obra* não aparece no texto original, pelo menos não aparece na parte em que estava disponível no *Compara*. No Quadro 20 aparecem alguns exemplos do item lexical *hearing* extraídos do *corpus Compara*.

Língua Inglesa (Texto A)	Língua Portuguesa (Texto B)
(1A) «Tubby, <u>I don't believe I'm hearing this</u> ,» said Ollie.	(1B) - Bolinha, <u>não acredito no que estou a ouvir</u> -- disse o Ollie.
(2A) As he went up he was already correcting the passage; <u>it</u>	(2B) Enquanto ia subindo, já estava a corrigir a passagem;

<u>was the French horns we should have been <b>hearing</b> now, not the clarinets, and the timpani markings were piano.</u>	<u>deviam ser as trompas a ouvir-se naquele momento, e não os clarinetes, e a marcação dos timbales devia ser piano.</u>
(3A) I've been <b>hearing</b> about your latest.	(3B) - Tenho ouvido falar da sua obra mais recente.

Quadro 20: Correspondentes Tradutórios do item lexical *hearing* + *verb to be* (Corpus Compara)

Outra co-ocorrência frequente, 8,88% no total do *corpus Compara*, foi o uso de *hearing* + *on*, ou vice versa, conforme apresentado no Quadro 21.

LINHAS DE CONCORDÂNCIA			
12	But when on first	<b>hearing</b>	of this arrangement I tried to inform Miss Kenton of it
13	Mr Kirwin, on	<b>hearing</b>	this evidence, desired that I should be taken into the room...
14	He had involuntarily shaken Salvador's hand on	<b>hearing</b>	his last words, then repented the impulsive gesture...
15	On	<b>hearing</b>	him, all of them felt a sudden and urgent desire to urinate,...
16	The girl with dark glasses who had got up on	<b>hearing</b>	his moan began approaching slowly, counting the beds.
17	...like a pet dog delirious with excitement on	<b>hearing</b>	its master return, all it had to do was to jump down...
18	...my reaction on receiving this book, my surprise on	<b>hearing</b>	that it carries no erratum, all of this is like having a sore,...
19	...calling to mind	<b>hearing</b>	on vaguely comparable

words he remembered	occasions, he announced,...
---------------------	-----------------------------

Quadro 21: Co-ocorrências do item lexical *on hearing* e *hearing on* (Corpus Compara).

No exemplo (12) do Quadro 21, *on first hearing* foi traduzido por *ao ter conhecimento*. O tradutor tentou adaptar o texto alvo para melhor entendimento dos leitores ao traduzir *hearing* por *conhecimento*. O exemplo (15), *on hearing him* foi traduzido por *ouvindo-o*, sendo que o verbo foi traduzido no gerúndio, e o item lexical *on* foi omitido ao ser traduzido. Já no exemplo (16), *on hearing* foi traduzido por *ao ouvir*, ficando o verbo no infinitivo. O tradutor utilizou uma expressão cultural ao traduzir *girl* por *rapariga* e omitiu a tradução de *dark* no texto em português, porém não houve perda na representação semântica do texto fonte para o texto alvo. No Quadro 22 apresentam-se os três exemplos citados acima em língua inglesa e seus respectivos correspondentes tradutórios em língua portuguesa.

<i>Língua Inglesa (Texto A)</i>	<i>Língua Portuguesa (Texto B)</i>
(1A) But when <u>on first hearing</u> of this arrangement I tried to inform Miss Kenton of it, she once again refused to converse with me, and in order to accomplish matters as quickly as possible I was actually obliged to write a note and put it under the door of her parlour.	(1B) Todavia, quando, <u>ao ter conhecimento</u> deste esquema, tentei informar Miss Kenton a esse respeito, ela recusou-se, mais uma vez, a falar comigo e tive realmente, para que tudo fosse preparado com a maior rapidez possível, de lhe escrever um bilhete e metê-lo por baixo da porta da sua sala.
(2A) <u>On hearing him</u> , all of them felt a sudden and urgent desire to urinate, and their thoughts were more or less as follows,	(2B) <u>Ouvindo-o</u> , todos sentiram uma súbita e urgente vontade de urinar, pensaram, por estas ou outras palavras,
(3A) The girl with dark glasses who had got up <u>on hearing</u> his	(3B) A rapariga dos óculos, que se tinha levantado <u>ao ouvir</u> o

moan began approaching slowly, counting the beds.	gemido, veio-se chegando devagar, contando as camas.
--	---

Quadro 22: Correspondentes Tradutórios do item lexical *on hearing* e *hearing on* (Corpus Compara).

No Quadro 23 aparecem duas ocorrências de *hard of hearing*. Nos dicionários Oxford (2000), Longman (2005) e Password (1998), encontraram-se a expressão *hard of hearing*, porém, não com as acepções de *hearing*, mas abaixo do item lexical *hard*. Nos dicionários Oxford (2000) e Longman (2005) há um link em *hearing* que leva a *see also HARD OF HEARING*.

LINHAS DE CONCORDÂNCIA			
20	Is it your impression that Philip Swallow is a bit hard of	hearing	?she asks him.
21	Perhaps I'm getting hard of	hearing	.

Quadro 23: Co-ocorrências do item lexical *hard of hearing* (Corpus Compara).

No exemplo (20) do Quadro 23, *hard of hearing* foi traduzido por *duro de ouvido*. Esse correspondente tradutório não soa natural em português do Brasil. Surpreendentemente, o dicionário Password (1998) traz como correspondente de *hard of hearing* a tradução *duro de ouvido*. Já o tradutor *Google Online*,<sup>39</sup> mostra como equivalente tradutório a palavra *surdo*. Ao fazer uma busca no *corpus* disponível do *Google*, o item lexical *surdo* apresenta uma frequência de ocorrências considerável (um milhão e noventa mil ocorrências), no entanto, *duro de surdo* não apresenta nenhuma ocorrência<sup>40</sup>.

O exemplo (21) *hard of hearing* foi traduzido por *endurecer os tímpanos*. Buscando por esta ocorrência entre *aspas* no *Google*, temos apenas 6 ocorrências de *endurecer os tímpanos*. A autora, não satisfeita

<sup>39</sup> Disponível em abril de 2010 < <http://translate.google.com.br/#en|pt|>>

<sup>40</sup> Disponível em julho de 2020 < <http://www.google.com.br/>>

com a resposta da pesquisa, buscou pelo *Google de Portugal*<sup>41</sup> para ver se havia alguma diferença, mas apareceram as mesmas 6 ocorrências anteriores. Isso significa que a escolha do correspondente tradutório não é muito frequente entre os falantes nativos de língua portuguesa. No Quadro 24, apresentam-se os correspondentes tradutórios dentro do contexto literário.

<i>Língua Inglesa (Texto A)</i>	<i>Língua Portuguesa (Texto B)</i>
(1A) 'Is it your impression that Philip Swallow is a bit hard of <b>hearing</b> ?' she asks him.	(1B)- Não tens a impressão de que o Philip Swallow é um bocadinho duro de ouvido? - pergunta-lhe Robyn.
(2A) Perhaps I'm getting hard of <b>hearing</b> ,	(2B) Talvez se me estejam a endurecer os tímpanos,

Quadro 24: Correspondentes Tradutórios do item lexical *hard of hearing* (Corpus Compara).

Na lista de concordância apresentada no Quadro 25, temos o item lexical *in + hearing*: *in his hearing, no difficulty in hearing, at least none in the hearing, in Rodney Wainwright's hearing, in the hearing of*.

<i>LINHAS DE CONCORDÂNCIA</i>	
22	as if it had been spoken in his <b>hearing</b> , and yesterday.
23	, for they had no difficulty in <b>hearing</b> all that was said on deck.
24	Peters or the cook -- at least none in the <b>hearing</b> of Augustus.
25	... the bus to return home he complains in Rodney Wainwright's <b>hearing</b> of feeling feverish, and spends the entire journey ...

<sup>41</sup> Disponível em abril de 2010 < <http://www.google.pt> >

26	<i>With the disturbance of the spring?» in the</i>	<b>hearing</b>	of the Americana's bell captain, to his considerable ...
----	--	----------------	---

Quadro 25: Co-ocorrências do item lexical *in hearing* (Corpus Compara).

No exemplo (22) do Quadro 25, *spoken in his hearing* foi traduzido por *na sua presença*. No Quadro 26, são mostrados os correspondentes tradutórios em que há omissão da palavra *spoken* e *in his hearing* foi traduzido por *na sua presença* o que não apresenta uma tradução literal do contexto. O item lexical *presença* não é um equivalente tradutório que podemos encontrar nos dicionários para *hearing*. No exemplo (23), a estrutura da frase foi mudada e transformada de frase negativa - *for they had no difficulty in hearing* - para frase afirmativa - *pois podiam ouvir perfeitamente*. Nesse contexto o item lexical *in* acaba sendo deixado do lado, pois o tradutor optou por mudar a estrutura da frase, contudo a representação semântica continua semelhante. No exemplo (24) do Quadro 25, *He complains in Rodney Wainwright's hearing of feeling feverish* o item lexical de pesquisa *hearing* foi traduzido pelo verbo *ouvir*.

<i>Língua Inglesa (Texto A)</i>	<i>Língua Portuguesa (Texto B)</i>
(1A) The biographer had been studying old court records and what had been said by the accused was fresh to him as if it had been spoken in his <b>hearing</b> , and yesterday.	(1B) O biógrafo tinha andado a analisar velhos autos judiciais e o que fora dito pelos acusados estava tão fresco na sua mente como se tivesse sido na sua presença e no dia anterior.
(2A) The mate gave them fair words -- no doubt with a view of inducing those below to yield, for they had no difficulty in <b>hearing</b> all that was said on deck.	(2B) O imediato dirigiu-lhes boas palavras, sem dúvida para levar os de baixo a submeterem-se, pois podiam ouvir perfeitamente o que se dizia na coberta.
(3A) As the party boards the bus to return home <u>he complains in Rodney Wainwright's <b>hearing</b> of feeling feverish</u> , and spends the entire jour-	(3B) Enquanto o grupo ia entrando para o autocarro, <u>Rodney Wainwright consegue ouvi-lo queixar-se de que se sente febril</u> , e passa toda a

ney with his eyes closed and an expression of suffering on his face, but joy is silent and unsympathetic, sitting beside him with her own eyes inscrutable behind dark glasses.	viagem com os olhos fechados e uma expressão de sofrimento estampada no rosto. Joy, no entanto, mantém-se calada e indiferente, sentada a seu lado, com os seus próprios olhos inescrutáveis por detrás dos óculos escuros.
---	---

Quadro 26: Correspondentes Tradutórios do item lexical *in hearing* (Corpus Compara).

Os exemplos listados acima representam um pequena amostra do item lexical *hearing* dentro do contexto literário (*Corpus Compara*). Não é intenção da autora mostrar todos os exemplos analisados, pois esses encontram-se em formato eletrônico para eventuais dúvidas.

### 3.3 Análise dos Resultados Gerais

Com base nos dados extraídos dos exemplos dos *corpora Euro-*parl** e *Compara*, compreende-se que as co-ocorrências do item lexical *hearing* são diferentes nos dois *corpora*. Observe alguns exemplos abaixo. Os itens que aparecem com (\*) são os que co-ocorrem nos dois contextos.

CO-OCORRÊNCIAS do item lexical “hearing” –EuroParl	CO-OCORRÊNCIAS do item lexical “hearing” -Compara
during the hearing last month	verb (to be) * I was hearing
after (the) hearing that statement about *	on hearing him *
held a hearing last March	in hearing all that was said
public hearing in October	He clothed himself upon hearing
hearing about assistance *	(pronoun- his/their/our/your) his hearing attuned to reading
look foward to hearing the Com- mission`s proposal	Hard of hearing
We keep on hearing/ * hearing on water policy	Of hearing *
(verb) to be + hearing*	From hearing*

## Quadro 27: Co-ocorrentes do EuroParl VS Compara

Ao analisar os colocados e coligações que co-ocorrem nos dois contextos percebeu-se que os padrões e correspondentes tradutórios eram diferentes, como segue abaixo no exemplo de *hearing* + verb to be no contexto literário versus legislativo.

O exemplo acima representa o contexto literário, o item lexical *hearing* foi traduzido pelo verbo *escutar*. Em inglês o tempo verbal apresentado foi o *past continuous* e ao ser traduzido para o português percebe-se que o tempo continuou o mesmo (o que chamamos de gerúndio).

I had just been through the Listening Exercise, but I couldn't believe what I was hearing from Petrus.

Apesar do exercício de ouvir, meus ouvidos se recusavam a acreditar o que eu estava escutando.

Já o exemplo abaixo, que representa o *corpus* legislativo, o item lexical *hearing* foi traduzido por *assistir*. Ao procurar no dicionário pelos correspondentes tradutórios de *hearing* não existe a opção de *assistir* como correspondente tradutório do item lexical *hearing*.

Os dois exemplos apresentados acima pretendem mostrar os padrões para cada contexto e evidenciar que as traduções ocorrem de forma diferente dependendo do contexto e colocados que acompanham a palavra.

Ao observar os correspondentes tradutórios no quadro abaixo, é notável que, no *corpus* jurídico/legislativo, os correspondentes de *hearing* são mais frequentemente traduzidos por *substantivos*, e, no *corpus*

... que no fundo veio para **ouvir** o Parlamento no seu todo , invertendo-se assim o verdadeiro sentido , pois não é já a comissão que assiste ao debate , mas sim , ao invés , é o Parlamento que ouve um parecer da comissão ; depois a comissão deixa o hemicíclio , já não participando no debate em si.

... who has actually come to hear Parliament , is given the floor , thereby turning the whole point of the exercise on its head , as the Commission is not hearing the debate but , quite the reverse , Parliament is hearing the Commission 's position . The Commissioner then departs and takes no further part in the actual debate .

literário, existe uma presença maior de *verbos*.

No *Compara*, o equivalente tradutório mais frequente é o verbo *ouvir* com 48,88% das ocorrências; 8,88% das ocorrências são do substantivo *ouvido*; 8,88%, do verbo *escutar*; 4,44% dos casos foram omisso; 2,22%, do item lexical *falar*; e 2,22%, dos substantivos *audição* (*aparelho auditivo*).

Os demais correspondentes tradutórios encontrados no *corpus Compara* tiveram ocorrência de apenas uma vez cada, somando um total de 11 ocorrências, caracterizando 10% no total do *corpus Compara*. Dentre os correspondentes encontrados, percebemos que *ter conhecimento* e *assistir* aparecem tanto no *corpus Compara* quanto no *Euro parl*. O termo *em sessão* que se refere à *audiência* apareceu uma vez no *Compara*, e a palavra *tímpano*, que é sinônimo de *ouvido*, segundo o dicionário *Priberam da Língua Portuguesa*<sup>42</sup>, também aparece no *Compara*.

No *corpus Euro parl*, os itens lexicais de maior frequência foram os substantivos *audição* e *audiência*, com 55,45% das ocorrências; o verbo *ouvir* aparece com 19,09% das ocorrências; os itens lexicais *assistir*, *consulta*, *exprimir*, *surdos*, *atenção*, *informação* e *falar* ocorreram duas vezes cada um (1,81%), e em 2,7% dos casos foram omitidos.

CORRESPONDENTES TRADUTÓRIOS – (Euro parl)	N.	CORRESPONDENTES TRADUTÓRIOS (Compara)	N.
Audição	52	Ouvir*	45
Ouvir *	21	Escutar*	8
Audiência	9	Ouvido*	8
Assistir *	2	Falar*	3
Consulta	2	Audição (ap. auditivo)	2
Exprimir	2	Saber	2
Surdos	2	Ter conhecimento*	1
Atenção	2	Vivendo	1
Informação	2	Pressentir	1
Falar*	2	Presença	1
Ter conhecimento*	1	Assistir *	1

Quadro 28: Correspondentes tradutórios do EuroParl VS Compara

<sup>42</sup> Disponível em maio de 2010 <<http://www.priberam.pt/dlpo/consultar.aspx>>

1. Cavidade do ouvido em que se estende a membrana sonora. 2. *Infrm*. Ouvido.

Percebe-se que, dependendo dos colocados ou coligações que acompanham um item lexical, e principalmente da especialidade do *corpus*, seu correspondente tradutório pode variar. Ao observar o *corpus Europarl*, o item lexical *during* acompanhado de *hearing* apresenta como correspondente tradutório o substantivo *audição*, como apresentado a seguir:

- a. ... já na sua *audição*, no mês passado
- b. ...lembro-me do senhor comissário Fischler afirmar durante a sua *audição*, perante a comissão...
- c. ...frisou aquando da sua *audição* na semana passada...
- d. ...no decorrer da *audição* pública em de Novembro passado...
- e. ...numa *audição* da Subcomissão...
- f. ...tendo dato, no decurso da sua *audição*, uma volta de 180 °...
- g. ...nesta altura foi decidida a realização da *audição* sobre os direitos do Homem...

Ainda no *corpus Europarl* percebe-se que *held* acompanhado de *hearing* também apresenta como correspondente tradutório o item lexical *audição*:

- a. ...realizamos uma *audição* em 27 de Janeiro de 1997...
- b. ...a comissão já efectuou uma *audiência* ao nível desta câmara ...
- c. ...realizamos no Parlamento, em março, uma *audição* com um membro da comissão...
- d. ...teve lugar em Bruxelas uma *audição* organizada pela subcomissão...

Ao buscar pelas *coligações* e *colocados* no *corpus Compara*, percebe-se que os correspondentes tradutórios que acompanham o item lexical *hearing* são parecidos. *On hearing* é mais traduzido pelo verbo *ouvir*, conforme demonstram estes exemplos:

- a. ... ao *ter conhecimento* deste esquema... (on *hearing*)
- b. ...depois de *ter ouvido* estas testemunhas... (on *hearing*)
- c. ...ao *escutar-lhe* as últimas palavras... (on *hearing*)
- d. ...*ouvindo-o*, todos sentiram uma súbita e urgente... (on *hearing*)
- e. ...que se tinha levantado ao *ouvir* o gemido... (on *hearing*)
- f. ...louco de entusiasmo ao *pressentir* o dono... (on *hearing*)
- g. ...a surpresa ao *ouvir* que não está emendado... (on *hearing*)

- h. ...recorrendo à lembrança doutras semelhanças *escutadas* em ocasiões mais ou menos parecidas... (*hearing on*)

Já em *hearing on* e *on hearing* do *corpus Europarl*, observa-se que nas letras de *c* a *f* a preposição *on* aparece depois da palavra *hearing* traduzido por *audição/audiência*. O *corpus* apresenta apenas dois exemplos com o verbo *ouvir* e nos dois casos *on* aparece na frente de *hearing*, como aparecem nos exemplos abaixo:

- a. ...constato, ao *ouvir* o que acaba de nos ler.. (*on hearing*)
- b. ...*ouve-se* constantemente dizer... (*on hearing*)
- c. ...convide estes peritos para uma *audição* sobre normas de segurança... (*a hearing on*)
- d. ...é uma *audição* para decidir a .... (*hearing on*)
- e. ...à *audição* do Parlamento sobre a política da água... (*hearing on*)
- f. ...por ocasião da *audiência* sobre esse tema... (*hearing on*)

Obviamente existe um contexto muito maior que irá determinar o equivalente de *hearing*, mas os *colocados* e *as coligações* mais próximos e, principalmente, a especialidade do *corpus* são necessários. As palavras se combinam e acabam assumindo um significado próprio, para este significado podemos usar palavras diferentes que expressem a mesma coisa, como neste exemplo: *I'm tired of hearing* tanto podemos dizer: *Estou cansado de ouvir*, *Custa-me ouvir* ou *Estou farto de ouvir*, em qualquer uma das situações estaria traduzindo *I'm tired of hearing*, que é uma expressão muito comum no inglês. Em português, seria normal falar para alguém: *Não vejo a hora de te ver novamente*, no entanto, não é comum dizer em inglês: *I don't see the time...* Nessa situação, o mais frequente seria dizer: *I can't wait to see you again*. A ilustração acima serve apenas para demonstrar que um bom dicionário deve apresentar exemplos reais dos léxicos apresentados, pois tanto para tradutores experientes, quanto para alunos que não têm muita experiência com o idioma, os *corpora* ajudam na escolha do melhor equivalente tradutório. Ao entrevistar professores, estes nos falam sobre uma infinidade de erros que os alunos cometem ao buscar no dicionário a tradução de determinado item lexical e usarem este item de forma indevida por falta de orientação no dicionário.

#### 4. CONSIDERAÇÕES FINAIS

O objetivo desta dissertação foi apresentar uma metodologia com base no uso de *corpora* como ferramenta de apoio para o tradutor profissional e para o aprendiz de tradução. Para demonstrar esse procedimento foram estudados os colocados, as coligações e as soluções tradutórias do item lexical *hearing* em dois *corpora* paralelos de áreas diferentes; o *Compara*, com textos literários; e o *EuroParl*, com texto da área jurídica/legislativa.

O presente trabalho salientou que o corpus pode favorecer tanto o tradutor experiente, auxiliando na visualização da padronização das frases e propondo sugestões de uso real da linguagem, quanto para o aprendiz de tradução, servindo como pesquisa complementar aos dicionários que normalmente não oferecem exemplos reais de uso. O *corpus* é uma ferramenta de apoio eficaz e necessária para o resultado de um texto de chegada mais fluente. Percebeu-se que os dicionários disponíveis atualmente no mercado não proporcionam uma fonte segura de correspondentes tradutórios e que não apresentam informações de uso.

A análise dos dois *corpora* de áreas diferentes nos ajudou a observar que o contexto e os co-ocorrentes influenciam na escolha do equivalente tradutório. Nos textos legislativos os termos mais frequentes foram mais técnicos, como, por exemplo: *audiência* e também mais formais como a expressão *look forward to hearing*, enquanto que nos textos literários os equivalentes encontrados foram palavras mais usuais do dia a dia, como o verbo *ouvir* ou o substantivo *ouvido*.

A pesquisa com base em *corpus* constitui uma pesquisa empírica em que os dados coletados representam uma fonte necessária e importante para os estudos e análises linguísticas. Para auxiliar essas análises linguísticas, o uso do *software WordSmith Tools* (Scott, 1999), é uma ferramenta digna de apreciação, pois podemos analisar grandes quantidades de informações. As concordâncias que podemos fazer com essa ferramenta *WordSmith* nos ajudam a esclarecer dúvidas em relação aos colocados e ao contexto em que a palavra se insere.

Esta metodologia de pesquisa, embora simples, pretende mostrar aos tradutores e pesquisadores não somente que os dicionários hoje existentes no mercado, apesar de bons, ainda precisam melhorar. Mas, que uma das fontes mais confiáveis de pesquisa é um *corpus* com fontes seguras. Para o tradutor nada melhor do que pesquisar a frequência que

determinados termos ou expressões na língua de chegada para são usados para criar um texto mais natural e fluente.

Para futuros pesquisadores fica a sugestão de explorar melhor o multilinguismo no *corpus Europarl*, pois este possui 23 línguas disponíveis gratuitamente na *web* para serem analisadas. A elaboração de glosários a partir do corpus é um exemplo. Outra pesquisa interessante a ser desenvolvida seria a continuação do presente trabalho em relação à prosódia semântica do item lexical de busca “*hearing*” que não foi contemplado neste trabalho.

## 5. REFERÊNCIAS

1. ARROJO, Rosemary. (1998). Os Estudos da Tradução como área de pesquisa independente: dilemas e ilusões de uma disciplina em (des)construção. Disponível em: <[http://www.scielo.br/scielo.php?pid=S0102-44501998000200007&script=sci\\_arttext&tlng=es](http://www.scielo.br/scielo.php?pid=S0102-44501998000200007&script=sci_arttext&tlng=es)> Acesso em: 17 agosto 2009.
2. BAKER, Mona. (1994). In Other Words. London & New York: Routledge.
3. BAKER, Mona. (1995). Corpora in Translation Studies: An Overview and some Suggestions for Future Research. Target 7.2, 223-243.
4. BÉJOINT, H. The foreign student's use of monolingual English dictionaries: a study of language needs and reference skills. Applied Linguistics, v. 2, n.3, p. 207-222, Autumn 1981.
5. BERBER Sardinha, A.P. (2002). Linguística de Corpus. São Paulo: Editora Manole.
6. BIBER, D.; Conrad, S.; Reppen, R. (1998). Corpus linguistics: investigating language structure and use. Cambridge, Cambridge University Press.
7. BOWKER, Lynn & Jeniffer Pearson. (2002). Working with Specialized Language – A practical guide to using corpora, London/New York: Routledge.
7. COMET. Projeto Comet: Corpus Multilíngue para o ensino de tradução. Disponível em <[www.fflch.usp.br/dlm/comet/consulta\\_cortrad.html](http://www.fflch.usp.br/dlm/comet/consulta_cortrad.html)> Acesso em: agosto 2009.
8. COMPARA. Disponível em: <<http://www.linguateca.pt/COMPARA/index.php>> Acesso em: agosto de 2008.
9. ENGWALL G. (1992). “Comments”. In: Svartvik, J. (ed). Direction in Corpus Linguistics: proceeding of Nobel Symposium 1992. Berlin/New York: Mouton de Gruyter. 164-169.
10. European Parliament Proceedings Parallel Corpus (1996-2006). disponível em: <<http://www.statmt.org/europarl/>> Acesso em: 18 agosto 2009.
11. FERNANDES, Lincoln. (2006). Corpora in Translation Studies: revisiting Baker's typology. Disponível em: <<http://www.periodicos.ufsc.br>> Acesso em: 18 agosto 2009.

12. FILLMORE, Charles J. (1992): “Corpus linguistics” vs. “computer-aided armchair linguistics”; in *Directions in Corpus Linguistics: Proceedings from a 1991 Nobel Symposium on Corpus Linguistics* (pp. 35-66), Mouton de Gruyter, Stockholm.
13. FIRTH, J.R. (1957). Papers in Linguistics. London. Oxford University Press.
14. FRANCIS, W.N. (1992). “Problems of assembling and computerizing large corpora”. In: Johansson, S. *Computer corpora in English language research*. 7-24. Bergen: Norwegian Computing Center for Humanities.
15. GOYOS JR, Durval de Noronha. *Noronha’s Legal Dictionary = Noronha Dicionário Jurídico*. 5ed. São Paulo: Observador Legal, 2003.
16. HUMBLÉ P. R. M. Dictionaries and Language Learners, Frankfurt am Main: Haag + Herchen Verlag, 2001.
18. HUNSTON, Susan. (2002). Corpora in Applied Linguistics. Cambridge: Cambridge University Press.
19. JOHANSSON, Stig. (1998). On the role of corpora in cross-linguistic research. In: Stig Johansson e Signe Oksefjell (eds). *Corpora and crosslinguistic research: teory, method and case studies*. Amsterdam: Rodopi.
20. KENNEDY, G. D. (1998). *An introduction to corpus linguistics*. Nova York, Longman.
21. KOEHN, Philipp. (2005). A Parallel Corpus for Statistical Machine Translation. Disponível em: <<http://www.iccs.inf.ed.ac.uk/~pkoehn/publications/europarl-mtsummit05.pdf>> Acesso em: 19 agosto 2009.
22. KOEHN, Philipp. EuroParl: A Parallel Corpus for Statistical Machine Translation. School of Informatics. University of Edinburgh, Scotland. Disponível em: <<http://www.iccs.inf.ed.ac.uk/~pkoehn/publications/europarl-mtsummit05.pdf>>. Acesso em: setembro de 2009.
23. LEECH, G. (1992). Corpora and Theories of linguistic performance. In: Jan Svartvik (ed) *Directions in corpus linguistics*, pp. 105-22. Berlin, Mouton de Gruyter.
24. LEECH, G. (1997). Teaching and Language Corpora: A convergence. In A. Wichmann et al. (eds). *Teaching and Language Corpora*. (pp. 1-24). Harlow: Longman.
25. MCNERY, A.M. & Xiao, R.Z. (2007). Parallel and Comparable Corpora: What are they up to? In: *Incorporating Corpora: Translation and the Linguistic*. Disponível em:

- <[http://eprints.lancs.ac.uk/59/1/corpora\\_and\\_translation.pdf](http://eprints.lancs.ac.uk/59/1/corpora_and_translation.pdf)> Acesso em 17 de agosto de 2009
26. MICHAELIS: minidicionário inglês. São Paulo. Editora Melhoramentos. 2008.
27. MINDT, D. (1994). English Corpus Linguistics and the Foreign Language Teaching. In: J. Thomas & M. Short (eds), *Using Corpora for Language Research*. London: Longman.
28. PAGANO, Adriana; VASCONCELLOS, Maria Lucia. (2003). Estudos da Tradução no Brasil: reflexões sobre teses e dissertações elaboradas por pesquisadores brasileiros nas décadas de 1980 a 1990. Disponível em: <[http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S0102-44502003000300003](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0102-44502003000300003)> Acesso em 17 de agosto de 2009.
30. PARTINGTON, Alan. (1998). Patterns and Meanings: Using Corpora for English Language Research and Teaching. Amsterdam: John Benjamins.
31. PASSWORD: English dictionary for speakers of Portuguese. 2ed; São Paulo: Martins Fontes, 1998.
32. ROCHA, Marco. (2003). O uso de corpora computadorizados em lexicografia: guia do consumidor. In: Dias da Silva, B.C. (org.). *Todas as trilhas: pesquisas e projetos*. Araraquara: Laboratório Editorial/FCL/UNESP; São Paulo: Cultura Acadêmica Editora, pp. 143-177.
33. SANTOS, Dianna. (1998). Perception verbs in English and Portuguese. In: Stig Johansson and Signe Oksefjell (eds.) *Corpora and Cross-linguistic Research: Theory, Method, and Case Studies*. Amsterdam/Atlant, GA: Rodopi, pp 319-342.
34. SARDINHA, Tony Berber. (2000). O que é um corpus representativo? Disponível em: <<http://www2.lael.pucsp.br/direct/DirectPapers44.pdf>> Acesso em: 17 agosto 2009.
35. SARDINHA, Tony Berber. (2004). Linguística de Corpus. Barueri, SP: Manole.
36. SARDINHA, Tony Berber. (2004). Linguística de Corpus: Histórico e Problemática. Disponível em: <[http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S0102-44502000000200005](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0102-44502000000200005)> Acesso em: 15 agosto 2009.
37. SINCLAIR, J. (1991). Corpus, concordance, collocation. Oxford: Oxford University Press.
38. SINCLAIR, J. (2004). How to Use Corpora in Language Teaching. Oxford: Oxford University Press

39. TAGNIN, S. (Org.) Cadernos de Tradução. Special volume on corpora and translation.
40. TAGNIN, Stella. (2002). Os Corpora: instrumentos de auto-ajuda para o tradutor. Cadernos de Tradução, Florianópolis, v.9, n. 2002/1, p.191-213, 2003.
41. TAGNIN, Stella. (2005). O jeito que a gente diz: expressões convencionais e idiomáticas inglês e português. São Paulo. Ed. Disal.
42. THUNES, Martha. (1998) Classifying translational correspondences. In: JOHANSSON, S.: OKSEFJELL, S. Corpora and Cross-linguistic Research. Amsterdam: Atlanta: GA.
43. TOGNINI-BENELLI, Elena. (2001). Corpus Linguistics at Work. Amsterdam: John Benjamins.
44. TURIGAL. Corpus: Português/inglês de turismo. Disponível em : <<http://sli.uvigo.es/CLUVI/>> Acesso em: agosto de 2009.
45. WATKINS, M. and SILVA, Cecília Mendes F.S. (2000). English Prepositions for Brazilians. Curitiba. Ed. UFPR.
46. WICHMANN, A. et al. (eds.). (1997). Teaching and Language Corpora. Harlow: Longman.
47. WILLIAMS, J. and CHESTERMAN, A. (2002). The Map – A Beginner’s Guide to doing research in Translation Studied. Manchester: St. Jerome.

## 6. ANEXOS

### ANEXO A – EuroParl - Os debates são apresentados por data.

The screenshot shows the European Parliament website interface. The main heading is 'Verbatim Report of Proceedings in original language Debates published in the Official Journal (by part-session)'. Below this, there are two sections for the years 1999 and 1998. Each section contains a table of sessions, with columns for the month and session numbers, and sub-columns for the locations Strasbourg and Brussels.

Year	Month	Strasbourg	Brussels
1999	May	03, 04, 05, 06, 07	
	April	12, 13, 14, 15, 16	
	March	08, 09, 10, 11, 12	22, 23
	February	08, 09, 10, 11, 12	24, 25
	January	11, 12, 13, 14, 15	27, 28
1998	December	14, 15, 16, 17, 18	02, 03

### Relato integral das sessões

- A compilação das intervenções proferidas pelos deputados em plenário é conhecida por Relato Integral das Sessões (RIS). Cada intervenção é transcrita na língua utilizada pelo orador, que poderá fazer correcções ao texto.

Esta versão, tradicionalmente denominada "Arco-íris", é publicada na Internet em formato PDF ao mesmo tempo que é enviada para a tipografia do Parlamento Europeu para a feitura da edição em papel.

Só é acessível por data - isto é, o dia da sessão.

- Alguns dias úteis mais tarde, está disponível uma versão em formato HTML. Esta versão é de consulta mais fácil e é, poucos

dias depois, integrada na "pesquisa por orador" ou "pesquisa por palavra no título".

- Esta edição multilingue é subsequentemente substituída por uma edição revista, que contém as eventuais correcções feitas pelos oradores.

- **Tradução provisória**

- Nas semanas seguintes à sessão, a versão "Arco-íris" é traduzida nas onze línguas oficiais da UE. Logo que as traduções são enviadas ao Parlamento Europeu pelas equipas de tradução freelance, são disponibilizadas ao público Internet em formato PDF.

- **Edição definitiva**

A versão final traduzida destas intervenções é publicada (em formato HTML) em Europarl - substituindo a edição provisória multilíngue - logo que foram efectuadas as correcções finais nas onze versões linguísticas.

**ANEXO B – DEBATES DO PARLAMENTO EUROPEU** – uma pequena amostra em 5 línguas das 11 traduzidas no parlamento (Português, Inglês, Espanhol, Francês e Italiano)

1. Debate do Parlamento Europeu em Português:

**Debates do Parlamento Europeu**  
**SESSÃO DE SEGUNDA-FEIRA, 3 DE MAIO DE 1999**

**2. Homenagem Póstuma**

**Presidente.** – Senhores Deputados, lamento ter de comunicar que, na semana passada, o nosso antigo colega, o Vice-presidente Hans Peters, faleceu. Muitos dos que aqui se encontram foram testemunhas da sua importante acção na Comissão do Emprego e dos Assuntos Sociais, que não era senão mais uma manifestação da sua constante dedicação aos temas sociais e sindicais que o trouxeram a esta assembleia.

Muitos de nós conhecemos também o especial empenho que pôs no tratamento das questões imobiliárias, no seio do grupo especializado da Mesa do nosso Parlamento.

Dedicou todo o seu esforço a esta assembleia durante muitos anos e, por isso, creio que, de acordo com a nossa tradição, posso solicitar-lhes que guardemos um minuto de silêncio em memória do nosso companheiro.

(O Parlamento guarda um minuto de silêncio)

2. Debate do Parlamento Europeu em Inglês:

**Debates of the European Parliament**  
**SITTING OF MONDAY, 3 MAY 1999**

**2. Tribute**

**President.** – Ladies and gentlemen, I am sorry to have to tell you that our former colleague and Vice-President Hans Peters died last week. Along with many other Members present here today, I remember his sterling work on Parliament's Committee on Social Affairs. This was just one of the ways in which he expressed his strong commitment to social and trade union affairs which led him to become a Member of this House.

Many of us will also recall the special interest he took in matters concerning premises as a member of the specialist group of Parliament's Bureau.

He gave his all to the House for many years. In accordance with our tradition, therefore, I feel it is appropriate to ask you to observe a minute's silence in memory of our colleague.

(Parliament observed a minute's silence)

## 3. Debate do Parlamento Europeu em Espanhol:

**Debates del Parlamento Europeo**  
**SESIÓN DEL LUNES, 3 DE MAYO DE 1999**

**2. Elogio póstumo**  
**El Presidente.** – Señorías, lamento tener que comunicarles que, la semana pasada, nuestro antiguo colega el vicepresidente Hans Peters falleció. Muchos de los que aquí estamos hemos sido testigos de su acción importante en la Comisión de Empleo y Asuntos Sociales de este Parlamento, que no era sino una manifestación más de su constante dedicación a los temas sociales y sindicales que le trajeron a esta Asamblea.

Muchos de nosotros también sabemos del especial empeño que puso en ocuparse de las cuestiones inmobiliarias en el seno del grupo especializado de la Mesa de nuestro Parlamento. Dedicó todo su esfuerzo a esta Asamblea durante muchos años y, por ello, creo que, de acuerdo con nuestra tradición, puedo pedir a ustedes un minuto de silencio en recuerdo de nuestro compañero.

(El Parlamento observa un minuto de silencio)

## 4. Debate do Parlamento Europeu em Francês

**Débats du Parlement européen**  
**SÉANCE DU LUNDI 3 MAI 1999**

**2. Éloge funèbre**  
**Le Président.** – Mesdames et Messieurs, je suis au regret de vous communiquer le décès de notre ancien collègue, le vice-président Hans Peter, la semaine dernière. Nombre d'entre nous ont été témoins de son importante activité au sein de la commission des affaires sociales du Parlement, signe parmi tant d'autres de l'intérêt constant qu'il manifestait pour les questions sociales et syndicales qui l'ont amené dans cette Assemblée.

Nombre d'entre nous se souviennent également que c'est avec opiniâtreté qu'il s'occupait des questions immobilières au sein du groupe spécialisé du Bureau de notre Parlement. Il a consacré tous ses efforts à cette Assemblée pendant de longues années et je pense dès lors que, conformément à la tradition de ce Parlement, je puis vous demander d'observer une minute de silence à la mémoire de notre collègue.

(L'Assemblée, debout, observe une minute de silence)

## 5. Debate do Parlamento Europeu em Italiano:

**Discussioni del Parlamento europeo****SEDUTA DI LUNEDI' 3 MAGGIO 1999****2. Necrologio**

**Presidente.** – Onorevoli deputati, ho purtroppo il triste compito di annunciarvi la scomparsa di un nostro ex collega, il vicepresidente Hans Peters, avvenuta la settimana scorsa. Molti di noi sono stati testimoni dell'importante attività che egli ha svolto in seno alla commissione per l'occupazione e gli affari sociali di questo Parlamento, dove ha dimostrato la sua incessante dedizione ai temi sociali e sindacali che lo avevano condotto in questa Assemblea.

Molti di noi hanno altresì constatato il suo particolare impegno a occuparsi delle questioni immobiliari all'interno del gruppo *ad hoc* dell'Ufficio di presidenza del nostro Parlamento.

Per anni ha lavorato instancabilmente e con totale dedizione per questa Assemblea e pertanto, come è nostra tradizione, vi chiedo di osservare un minuto di silenzio in memoria del nostro collega.

(Il Parlamento osserva un minuto di silenzio)