

## **Analysing the product of translation with corpora**

**Silvana Maria DE JESUS**  
**Universidade Federal de Minas Gerais**

### **Como citar este artículo:**

DE JESUS, Silvana Maria (2005) «Analysing the product of translation with corpora», en ROMANA GARCÍA, María Luisa [ed.] *II AIETI. Actas del II Congreso Internacional de la Asociación Ibérica de Estudios de Traducción e Interpretación. Madrid, 9-11 de febrero de 2005*. Madrid: AIETI, pp. 588-601. ISBN 84-8468-151-3. Versión electrónica disponible en la web de la AIETI:  
<[http://www.aieti.eu/pubs/actas/II/AIETI\\_2\\_AFG\\_Claves.pdf](http://www.aieti.eu/pubs/actas/II/AIETI_2_AFG_Claves.pdf)>.



# ANALYSING THE PRODUCT OF TRANSLATION WITH CORPORA

Silvana Maria de Jesus  
Universidade Federal de Minas Gerais<sup>1</sup>

## 1 Introduction

This paper aims to present the results of a short research on reporting verbs developed under the field of corpus-based translation studies. The Halliday's systemic functional grammar (1994) and the theoretical-methodological approach of Corpus Linguistics are adopted as the theoretical ground.

This work focuses on describing the differences and similarities in the way the reporting verbs THINK and PENSAR<sup>2</sup> are used in a novel in English and in its Brazilian-Portuguese translation. The novels are *Point counter point* by Aldous Huxley (1994), first published in 1928, and its translation, *Contraponto* (Huxley, 1971) done by Érico Veríssimo in 1934. This corpus is part of CORDIALL (Corpus of Discourse for the Analysis of Language and Literature), a project developed at NET (Núcleo de Estudos da Tradução da FALE/UFMG), which has been developing discursive corpus-based translation studies.

## 2 Theoretical and methodological basis

### 2.1 Corpus-based Translation Studies

According to Berber Sardinha (2000), the first corpus – the Brown corpus – came into sight in a time when Chomsky's ideas were the main linguistic trend, i.e., in a time when real occurrences of language were not investigated because the examples the linguistics needed were mostly made up. However, in the 1970s, new linguistic theories, such as “the social interpretation of language and meaning” by Halliday (1978) introduced the need of investigating real language events; this was the beginning of Corpus Linguistics.

Ever since, many corpora have been built, especially in England and U.S.A, not only because of linguistic interest in describing language, but also for translating and educational

---

<sup>1</sup> PhD student at Universidade Federal de Minas Gerais, Brazil, doing research period at Universidad Autónoma de Barcelona with a scholarship from CAPES - Coordenação de Aperfeiçoamento de Pessoal de Nível Superior.

purposes. Corpus (corpora = pl.) is defined by Baker (1995:225) as “a collection of texts held in machine-readable form and capable of being analysed automatically or semi-automatically in a variety of ways”.

Baker’s classification of types of corpora became seminal in this field. Baker (1995:230-235) suggests three kinds of corpora: i) parallel corpora: “source language-texts in language A and their translated versions in language B”; ii) multilingual corpora: “sets of two or more monolingual corpora in different languages”; and iii) comparable corpora: “two separate collections of texts in the same language” - original and translated texts. Parallel and comparable corpora are more interesting for translation. In this research, a small parallel corpus is used.

Another important aspect related to corpus is size. At first, attention was directed at large corpora (more than one million words), but lately, several translation researchers have developed discursive analyses based on small corpora, such as Ghadessy & Gao (2001), Munday (1998, 2002), Cruz (2003), Jesus (2004) and Pagano (2004). According to Sinclair (2001:x), more important than size is the methodological approach used by the researcher. Thus, Sinclair defines small corpora as the ones “designed for early human intervention (EHI) while large corpora are designed for late or delayed human intervention (DHI)”. Halliday’s (1994) systemic functional grammar has been used by translation scholars as a productive tool for discursive analyses in small corpora.

## **2.2 Halliday and his Systemic Functional Grammar**

Halliday’s (1994:xiii) theory considers that language can be classified in three metafunctions: ideational, interpersonal and textual. These metafunctions are:

the manifestations in the linguistic system of the two very general purposes which underlie all uses of language: (i) to understand the environment (ideational), and (ii) to act on the others in it (interpersonal). Combined with these is a third metafunctional component, the ‘textual’, which breathes relevance into the other two.

Halliday (2001:16) considers that “as a general rule, ‘translation equivalence’ is defined in ideational terms; if a text does not match its source text ideationally, it does not qualify as a translation”. Halliday (1994:106) considers that our experience is represented by language

---

<sup>2</sup> I am using CAPITAL LETTERS to represent lemmas: “a lemma is a label under which all the inflected forms of a word can be gathered” (Kenny, 2001:34).

through the “transitive system” or “process types”, which could be classified in six groups: existential processes (about happening), material processes (about doing), behavioural processes (about acting), mental processes (about sensing), verbal processes (about meaning), and relational processes (about being). In addition to the process, it is important to consider the participants and the circumstances involved in it.

This analysis focuses on mental process, more specifically, on how the thoughts of the characters are presented in the corpus, considering that *Point counter point* is a novel of ideas or a “roman à thèse” (Lodge, 1992:198), i.e., a work where the ideas of the characters are more important than the plot itself. Halliday (1994:118) classifies the mental processes into three sub-types: i) perception (seeing, hearing), ii) affection (liking, fearing), and iii) cognition (thinking, knowing). In order to analyse how the thoughts of the characters are presented, this investigation focuses on the mental processes of cognition THINK and PENSAR. These processes were chosen because THINK is very productive in English novels to represent people’s thoughts, ideas and opinions and PENSAR is its more expected correspondent in Portuguese.

Two participants are involved in the process of thinking. Halliday (1994:118) considers that mental processes are always transitive, because they “potentially involve both a Senser and a Phenomenon”. Usually, senser and process are in one clause (the reporting clause) and the phenomenon in another one (the reported clause). These examples are from the corpus of this research and show the mental process, the senser and the phenomenon:

- 482 'What a fool!' <**phenomenon**> she <**senser**> had thought <**process**>  
 454 he <**senser**> thought <**process**> it was one of these ridiculous fusses about nothing  
 <**phenomenon**>  
 496 I <**senser**> think <**process**> of you so incessantly <**phenomenon**>

In these examples, different types of phenomenon are shown. The types of phenomenon are relevant in this analysis because they are related to the way the thoughts of the characters are represented. Halliday (1994:115) considers that the phenomenon is **projected** by the mental process.

Halliday (1994:250) considers that **projection** is “the logical-semantic relationship whereby a clause comes to function not as a direct representation of (non-linguistic) experience but as a representation of a (linguistic) representation”. He points out that this

interdependency between the reporting and the reported clause can happen in three ways: i) **parataxis** – the two clauses are independent; ii) **hypotaxis** – the two clauses are dependent; and iii) **embedding** – “a clause or phrase comes to function as a constituent within the structure of a group” (Halliday, 1994:242). Still, according to Halliday (1994:252), these three types of interdependency between clauses are directly connected with the three kinds of projection<sup>3</sup>:

i) **Quoting** – two clauses are connected in an independent way (parataxis); the reporting clause projects the reported one as a wording or a quote, such as in *Mary said: “I will come back here to-morrow”*;

ii) **Reporting** – two clauses are connected in a dependent way (hypotaxis); the reporting clause projects the reported one as meaning, such as in *Mary thought she would go back there the next day*;

iii) **Embedding** – where a clause is projected by a verbal process noun, such as in *the thought that she might one day be a queen*.

It is relevant to point out that these types are not isolated categories; on the contrary, they are more like a continuum along with many others categories that can be included.

According to the categories proposed, five kinds of phenomenon are investigated in this paper, related to the mental processes of cognition THINK and PENSAR<sup>4</sup>:

i) **discontinuous report**<sup>5</sup> – when the phenomenon is implicit in the context or explicit in another sentence, or it is not mentioned. E.g. *Well, you thought wrong*;

ii) **thing** – when the phenomenon is not a clause, but a nominal group or a prepositional phrase. E.g. *He thought of his brother Tom* ;

iii) **embedding** – when the projection is done by a noun. E.g. *a chronic annoyance at the thought that all the Old Man's intellectual and moral virtues...*

iv) **paraphrase** – when the reported clause is represented as meaning. E.g. *I thought we'd agreed not to talk about it again*.

v) **quote** – when the reported clause is represented as wording. E.g. *“Like a force of nature.” he thought as he watched her*.

---

<sup>3</sup> Exemplos from Halliday, 1994:252.

<sup>4</sup> Exemplos from the corpus.

<sup>5</sup> This term is used by Thompson (1994:27) to refer to cases that “the reporting signal is carried by a grammatically separate structure”. Halliday (1994:81) refers to “discontinuous residue”.

These five types are associated with the four functions that the phenomenon can have according to Thompson's work presented in the next section. Thompson's theory is very close to Halliday's. Thus, in this research, they will be used together.

### 2.3 Thompson's "Reporting Guide"

Reporting is a very common function of language and it is used to report people's speech, thoughts and perceptions. Traditional English grammars are very restricted when presenting possible functions and structures of reporting. Usually, they stick to reported speech only, presenting the two more common structures: direct and indirect speech. However, after the spread of corpora in the 1990s, new grammars and dictionaries, based on large banks of real communicative events, have been showing how complex and productive reporting structures are in English.

Thompson (1994), in *Collins Cobuild English Guides 5: Reporting*, which is based on The Bank of English (corpora that contains over 200 million words), provides a helpful classification of many kinds of reporting structures and their functions. He also considers the different uses of reporting speech: in journalism, conversation, novels and academic writing. The structures presented by this author are mainly applied to reporting speech, but they can also be used for reporting thoughts and perceptions. According to Thompson (1994:120), we report thoughts as if we could listen the words "passing through someone's mind".

Besides classifying the types of phenomenon, Thompson (1994:84) associates them to four functions: i) reporting **words**; ii) reporting **meaning**; iii) given a **general idea** of the message; and iv) given **no message**. Table 1 shows the relation between the five kinds of phenomenon and the four functions they can perform (the examples are from the corpus):

Table 1 - Functions of the phenomenon

Function	Phenomenon's structure	Example
word	Quote	'And what about me?' he suddenly <b>thought</b>
meaning	paraphrase	He didn't <b>think</b> she would
	embedding	He agonized himself with the <b>thought</b> that he could never, never ask Susan
general idea	thing (prepositional phrase or nominal group)	He <b>thought</b> of Carling 'I don't <b>think</b> anything
no message	Discontinuous	His <b>thoughts</b> returned to Marjorie

This research presents an analysis of THINK in the original and of PENSAR in the translation, focusing on what kind of phenomenon and what kinds of function are associated to these reporting verbs. Next section presents a brief explanation on the software used – the *WordSmith Tools*, as well as the methodology used for the analysis

## 2.4 WordSmith Tools

*WordSmith Tools* was first published in 1996 by Oxford University Press. It was designed by Mike Scott, and it is an integrated suite of tools for investigating language. With *WordSmith Tools* it is possible to look at words in context or isolated.

There are three main tools: WordList, Concord and KeyWords; and three main utilities: Splitter, Text Converter and Viewer. This paper briefly presents the ones that are used in this research: WordList and Concord. Further information on *WordSmith Tools* is available on Berber Sardinha (1999, 2001) and Scott (1999, 2001).

WordList tool shows a list of all the words of a text, both in alphabetical and frequency order. It also gives statistical data about the texts used, such as their size in bytes, the number of different lexical items (types) and the total number of words (token), and other statistical numbers.

Concord tool shows a word or phrase within a context, by looking for a specific searching word or node in the selected corpus. The searching word is presented in *concordance lines* that allow you to see the co-text and word-clusters around your searching word. This information can be displayed easily in many ways, either alphabetical or file order.

In the next section, the corpus of this research is presented and analysed using *WordSmith Tools*. This software has been used productively in corpus based translation studies. It has showed great potential for automatic treatment of electronic texts in Corpus Linguistic.

## 2.5 Using WordSmith Tool to analyse THINK and PENSAR in the research corpus

As previously mentioned, the corpus of this research is compounded of the English novel *Point counter point* by Aldous Huxley (1994), first published in 1928, and of its Brazilian-Portuguese translation, *Contraponto*, by Érico Veríssimo (Huxley, 1971). Table 2 **Table 2**

shows the corpus' s statistical data given by WordList tool. The translation is bigger than the original in bytes and tokens, an important fact when we consider the frequency of the verb THINK and PENSAR in both texts. The translation also has a higher number of types, but these data will not be relevant for the scope of this research.

Table 2 – Statistical data of the corpus

Characteristics/ Novels	<i>PCP</i>	<i>CP</i>
Bytes	1.005.57	1.132.804
Types	14.407	19.490
Tokens	172.634	186.253

*PCP: Point counter point, CP: Contraponto*

After checking with the WordList tool the basic statistical data of the corpus, this tool was used to verify the forms and frequency of THINK/PENSAR. This can be done by using the alphabetical wordlist (the numbers in parentheses are the total occurrences for each form in the corpus):

*Point counter point:*

think (196), thinker (2), thinkers (1), thinking (99), thinks (6), thought (179), thoughtful (1), thoughtfulness (1), thoughts (38).

*Contraponto:*

pensa (19), pensadas (1), pensando (6), pensador (4), pensadores (1), pensam (1), pensamento (22), pensamentos (45), pensamos (3), pensando (33), pensar (61), pensara (3), pensarão (1), pensarem (1), pensariam (1), pensarmos (1), pensas (7), pensaste (2), pensativa (3), pensativo (2), pensava (69), pensavam (1), pense (7), pensée (1). pensei (8), pensem (1), pensemos (3), penses (2), penso (14), pensou (89).

Using the button *join* in the alphabetical wordlist, it is possible to add all these forms, obtaining the total number of occurrences: THINK (496) and PENSAR (412). It is interesting to note that there are 9 forms of THINK and 30 of PENSAR. It is quite natural since Portuguese verbs present more conjugated forms than English verbs. The total number of occurrences (THINK=496, PENSAR=412) shows a higher number of the mental verb in the original.

The following step is using Concord to check the concordance lines and the co-text of occurrence of the search words, THINK and PENSAR. The search word in English is

*think\*/thought\**, which means that Concord will look for any word which starts with *think* and *thought*, such as *thinking* and *thoughts*, at the same time. The result was 496 concordance lines for the original. The search word in Portuguese is *pens\**, which means now that Concord will look for any word with the radical *pens*. The result was 414 concordance lines for the translation. These are the first 3 lines as examples, arranged according to the order they appear in the novels :

*Point counter point*

N Concordance  
 1 oo certainly.' For of course, he was **thinking** (with Lucy Tantamount's im  
 2 ing and enjoying, loving and hating, **thinking**, remembering, imagining. A  
 3 nes. He smiled. But if only, he was **thinking**, she would leave me in pe

*Contraponto*

N Concordance  
 1 a.Porque, sem dúvida nenhuma, **pensava** Walter, assediado pela inex  
 2 sofre e goza, que ama e odeia, que **pensa**, que recorda, que imagina. E  
 3 dera que ela me deixasse em paz!", **pensou** consigo. Franziu os lábios

These lines are used to observe the co-text of THINK and PENSAR and they can be grown in order to present more co-text, if necessary. Through these concordance lines the kinds of phenomenon that THINK and PENSAR project can be observed. In order to analyse the phenomenon it is necessary to do some annotation. According to Hunston (2002:79) “corpus annotation is the process of adding information to a corpus”. The information that will be added is the classification of the types of phenomenon of THINK and PENSAR, according to the five types of phenomenon indicated: quote, paraphrase, embedding, thing and discontinuous. These examples show some corpus’ s annotation:

*Point counter point*

N Concordance  
 1 oo certainly.' For of course, he was <paraphrase> **thinking** (with Lucy Tantamount's im  
 2 g and enjoying, loving and hating, <discontinuous> **thinking**, remembering, imagining. A  
 3 nes. He smiled. But if only, he was <quote> **thinking**, she would leave me in pe

*Contraponto*

N Concordance  
 1 Porque, sem dúvida nenhuma, <paraphrase> **pensava** Walter, assediado pela inexor  
 2 goza, que ama e odeia, que <discontinuous> **pensa**, que recorda, que imagina. E o  
 3 e dera que ela me deixasse em paz!", <quote> **pensou** consigo. Franziu os lábios e a

The Concord tool allows arranging the concordance lines in many ways. In the examples above they are arranged according to the order THINK and PENSAR appear in the files.

### 3 Reporting what characters think

#### 3.1 Types of phenomenon and their functions

As previously mentioned, there are 496 occurrences of THINK in the original and 412 occurrences of PENSAR in the translation. The analysis of the concordance lines eliminated some of them, either because they are from a different semantic field, such as *pensão* (a kind of hotel) or because the number of occurrences is not significant. For example, the annotation shows that the occurrence of the structure embedding is not significant in the corpus (only 4 in the original and 2 in the translation), so this structure was eliminated, because they are not representative. The final numbers are: THINK (487) and PENSAR (398), and they are analysed according to four types of phenomenon: quote, paraphrase, thing and discontinuous.

Table 3 shows the occurrence of the four types of phenomenon in both novels. As already mentioned, the four types of phenomenon are directly associated with the functions of the message.

Table 3 - Types and functions of the phenomena in the novels

Function \ Category	<i>Point counter point</i>		<i>Contraponto</i>	
	Occurrence	%	Occurrence	%
Word / quote	103	21	124	31
Meaning / paraphrase	139	29	42	11
General idea / thing	149	31	128	32
No message / discontinuous	96	20	104	26
TOTAL	487	100	398	100

The data show that quote, thing and discontinuous reporting are quite similar in the original and in the translation, although it shows some variation. Paraphrase, on the other hand, varies a lot. It occurs only in 11% of the occurrences in the translation while it occurs in 29% of the original.

**Table** Table 4 shows occurrences of the translation that keep the same type of phenomenon as in the original and **Table** Table 5 shows examples of different types of phenomenon in the translation.

Table 4 – THINK/PENSAR – the same type of phenomenon in original and translation

Type	<i>Point counter point</i>	<i>Contraponto</i>
Quote	12 'And what about me?' he suddenly <b>thought</b> .	14 "E eu?", <b>pensou</b> ele subitamente.
Paraphrase	81 Mr Sita Ram, however, <b>thought</b> that the only hope was...	78 O Sr. Sita Ram, no entanto, <b>pensava</b> que a única esperança era...
Thing	11 He <b>thought</b> of Carling.	13 <b>Pensou</b> em Carling.
Discontinuous	4 He felt ashamed of his odious <b>thoughts</b> .	4 Walter sentiu-se envergonhado de seus <b>pensamentos</b> odiosos.

Table 5 – THINK/PENSAR different type of phenomenon in original and translation

<i>Point counter point</i>	<i>Contraponto</i>
36 is very unlikely, when you come to <thing> <b>think</b> of it. Very mysterious and...' Hi	41 a definida é coisa muito improvável, se <discontinuous> <b>pensarmos</b> bem. Muito misterioso e...
58 'nly I want a revolution.' The Old Man <thing> <b>thought</b> in terms of geology and was	62 dúvida: eu quero uma revolução. - Os <discontinuous> <b>pensamentos</b> do velho estavam dentro
67 pered by a chronic annoyance at the <embedding> <b>thought</b> that all the Old Man's intelle	63 contrariedade crônica que lhe vinha de <paraphrase> <b>pensar</b> que todas as virtudes intelectu

Observing the concordance lines classified as paraphrase in the original, it is possible to notice that in most of these lines THINK is used to express the opinions of the characters. This meaning of THINK corresponds in Portuguese to ACHAR. This research does not analyse the verb ACHAR, but a quick look into the concordance lines shows that many occurrences of THINK were translated as ACHAR. There are some examples in Table 6:

Table 6 – THINK/ACHAR

THINK expressing opinion in the original		ACHAR expressing opinion in the translation	
N	Concordance	N	Concordance
6	'Yes,' he said, 'I <b>think</b> that's true.	6	Sim - disse -, eu <b>acho</b> que é verdade.
13	'I <b>think</b> it's time,' Spandrell	15	- <b>Acho</b> que já é tempo
14	'I <b>think</b> I'll take to science, lik	16	- <b>Acho</b> que vou me dedicar à ciência
17	'I <b>think</b> you're a darling,' said	19	- <b>Acho</b> que tu és um querido -
20	I <b>think</b> we ought to buy some	25	<b>Acho</b> que devíamos comprar
22	'I <b>think</b> I shall run up to town	29	- <b>Acho</b> que vou à cidade
33	'I <b>think</b> I'll lie down for a moment	35	- <b>Acho</b> que vou deitar-me
34	. 'I <b>think</b> he ought to go to sleep	38	- <b>Acho</b> bom ele dormir.

Thompson (1994:125) points out that *think* is the most common verb to report someone's thought, such as in *She thought I was joking*. He also shows that *think* can be used to express your own opinion, which Thompson (1994:138) does not consider as real reporting. In this paper, *think* is analyzed as a reporting verb in a broad sense, so the phenomenon of this kind is also considered a paraphrase. Hence, the translation presents fewer occurrences of paraphrase since *think* for personal opinion was translated as *achar*, its more common correspondent in Portuguese.

The analysis of the kinds of phenomenon indicates the use of thing (nominal group and prepositional phrase) as the most frequent pattern in both novels, showing that about 30% of the thoughts of the characters are not fully presented, but just summarized. The second most frequent pattern for the original is paraphrase, i.e., thoughts are presented as meaning; and for the translation, the second most frequent pattern is quote, which means that the translation tends to present thoughts more as wording. An analysis of the verb ACHAR would be necessary in order to complete the analysis of paraphrase in the translation.

## 4 Conclusion

In order to analyse how the thoughts of the characters were represented in the English novel, *Point counter point*, and in its Brazilian-Portuguese translation, *Contraponto*, this research investigated the reporting verbs THINK and PENSAR and the kind of phenomenon these verbs project.

The analysis of what kind of phenomenon is used with these verbs shows that, both in the original and in the translation, the most frequent kind of phenomenon is what Halliday

calls thing, i.e., a nominal group or a prepositional phrase, which gives a general idea of the thought of the character.

This result shows that, although most grammars books (Murphy, 1994:94, Brinton, 2000:217) present quote (direct discourse) and paraphrase (indirect discourse) as the main patterns, the use of thing (prepositional phrase/nominal group) is a common pattern, being the main pattern of both novels herein analysed. This may mean that this category needs more attention from linguistic researches. In fact, all types of phenomenon are almost as productive as quote and paraphrase in the corpus.

Anotne important point of this research was to demonstrate how to use *WordSmith Tools* for productive linguistic research, especially the use of the utility Viewer, showing how rich and fruitful the approach of corpus-based translation studies is.

## 5 Bibliography

BAKER, M. Corpora in translation studies: an overview and some suggestions for future research. *Target*, Amsterdam, v. 7, n. 2, 1995: 223-243.

BERBER SARDINHA, Tony. Usando WordSmith Tools na investigação da linguagem. *DIRECT Papers 40*, 1999. Disponível em: < <http://sites.uol.com.br/tony4/homepage.html> >. Acesso em: 12 março 2003.

BERBER SARDINHA, Tony. *Linguística de Corpus: Histórico e Problemática*. DELTA. [online]. 2000, vol.16, no.2 [cited 16 April 2003], p.323-367. Available from World Wide Web: <http://lael.pucsp.br/~tony/download.htm>

BERBER SARDINHA, Tony. Comparing Corpora with WordSmith Keywords. *The ESPecialist*, São Paulo, v. 22, n. 1, p. 87-99, 2001.

BRINTON, Laurel J. *The structure of modern English: a linguistic introduction*. Amsterdam: John Benjamins Publishing Company, 2000.

CRUZ, Osilene Maria de Sá e Silva da. “*Harry Potter and the chamber of secrets*” e sua tradução para o português do Brasil: uma análise dos verbos de elocução com base na linguística sistêmica e nos estudos de corpora. Belo Horizonte: Faculdade de Letras da UFMG/PosLin, 2003. (Dissertação, Mestrado em Letras/Linguística, inédita).

GHADESSY, M. et al. (eds.). *Small corpus studies and ELT. Theory and practice*. Amsterdam: John Benjamins, 2001.

GHADESSY, M., GAO, Y. Small corpora and translation. Comparing thematic organization in two languages. In: GHADESSY, M. et al. (eds.). *Small corpus studies and ELT. Theory and practice*. Amsterdam: John Benjamins, 2001: 335-359.

HALLIDAY, M.A.K. *Language as social semiotic: the social interpretation of language and meaning*. London: Arnold, 1978.

HALLIDAY, M.A.K. *An introduction to functional grammar*. 2<sup>nd</sup> ed. London: Arnold, 1994.

HALLIDAY, M.A.K. Towards a theory of good translation. In: STEINER, E., YALLOP, C. (ed.). *Exploring translation and multilingual text production: beyond content*. Berlin & NY: Mouton de Gruyter, 2001: 13-18.

HUNSTON, Susan. *Corpora in applied linguistics*. Cambridge: Cambridge University Press, 2002: 38-66.

HUXLEY, Aldous. *Point counter point*. London: Flamingo, 1994.

HUXLEY, Aldous. *Contraponto*. Trad. Érico Veríssimo. Porto Alegre: Editora Globo, 1971. (Tradução de: *Point counter point*).

JESUS, Silvana Maria de. *Representação do discurso e tradução: padrões de textualização em corpora paralelo e comparável*. Belo Horizonte: Faculdade de Letras, UFMG/PosLin, 2004. (Dissertação, Mestrado em Letras/Linguística Aplicada).

KENNY, D. *Lexis and creativity in translation: a corpus-based study*. Manchester, UK & Northampton MA: St Jerome Publishing, 2001: 22-72.

LODGE, David. *The Art of Fiction*. NY & London: Penguin Books, 1992: 197-200.

MUNDAY, Jeremy. Problems of applying thematic analysis to translation between Spanish and English. *Cadernos de Tradução* III. Florianópolis: UFSC, 1998: 183-213.

MUNDAY, Jeremy. Systems in Translation. A systemic model for descriptive translation studies. In: Hermans, Theo (Ed.). *Crosscultural transgressions*. Research Models in Translation Studies II: Historical and Ideological Issues. Historical and Ideological Issues. Manchester: St. Jerome, 2002.

MURPHY, Raymond. *English grammar in use*. 2<sup>nd</sup>. ed. Cambridge: Cambridge University Press, 1994.

PAGANO, Adriana S. Organização temática e tradução. In: ALVES, Fabio, MAGALHÃES, Célia e PAGANO, Adriana (Ed.). *Competência em tradução: cognição e discurso*. Belo Horizonte: Editora da UFMG, 2004.

SCOTT, Mike. Comparing corpora and identifying key words, collocations, frequency distributions through the WordSmith Tools suite of computer programs. In: GHADESSY, M.

et al. (eds.). *Small corpus studies and ELT. Theory and practice*. Amsterdam: John Benjamins, 2001: 47-67.

SCOTT, Mike. *WordSmith Tools*. Oxford University Press, 1999. Disponível em <<http://www.lexically.net/wordsmith/>> Acesso em: 12 março 2003.

SINCLAIR, J.M. Preface. In: GHADESSY, M. et al. (eds.). *Small corpus studies and ELT. Theory and practice*. Amsterdam: John Benjamins, 2001: vii-xv.

THOMPSON, Geoff. *Collins Cobuild English Guides 5: Reporting*. London: HarperCollins Publishers, 1994: v.32, 120-135.