

Describing the methodological procedure of a translation studies investigation based on corpora

Sinara de Oliveira Branco*

Resumo: Este artigo tem o objetivo de apresentar os procedimentos metodológicos utilizados para o desenvolvimento de um estudo na área de Estudos da Tradução baseado em um Corpus Linguístico. Primeiramente, serão apresentadas as visões de teóricos que tratam do uso dos Corpora aplicados aos Estudos da Tradução, bem como as Estratégias de Tradução de Chesterman (1997). O corpus selecionado para a pesquisa é uma compilação de reportagens jornalísticas *online* dos sites Reuters, BBC, *The New York Times on the Web* e UOL *Jornais* sobre conflitos no Oriente Médio, originalmente escritos em inglês e traduzidos para o português brasileiro. Serão observadas, nas reportagens, as estratégias de tradução utilizadas seguindo o quadro de Estratégias de Tradução sugerido por Chesterman (1997) a partir dos critérios sintáticos, semânticos e pragmáticos, com o intuito de refletir sobre as práticas de tradução relacionadas a reportagens jornalísticas *online*. A metodologia da pesquisa será descrita para que, em seguida, sejam apresentados os procedimentos de análise do corpus. Após a análise, serão feitas considerações sobre o trabalho de tradução de textos jornalísticos *online* seguidos de comentários sobre a contribuição dos Corpora nos Estudos da Tradução.

Palavras-chave: Estudos da Tradução; Estudos Baseados em Corpora; Estratégias de Tradução; Tradução e Jornalismo *Online*.

Abstract This paper aims at presenting the methodological procedure used in order to carry out a study in the area of Translation Studies based on Corpus Linguistics (CL). Theoretical points of view regarding Translation Studies based on Corpora will be presented, as well as Chesterman's (1997) Translation Strategies. The corpus selected for the study is a compilation of online news reports from Reuters, BBC, The New York Times on the Web and UOL *Jornais* regarding Middle East conflicts originally written in English and translated into Brazilian Portuguese. The online news reports will be analysed in order to identify the translation strategies suggested by Chesterman (1997) following the criteria – from syntactic to semantic and then to pragmatic – in order to reflect on the translation practice concerning online news reports. The research methodology will be described followed by the corpus analysis description. After the corpus analysis, some comments on the translation of online news reports as well as on the contribution of Corpora to Translation Studies will be made.

Keywords: Translation Studies; Corpus-Based Studies; Translation Strategies; Translation and Online Journalism.

1 Translation and Online Journalism

In the year 2003, two Brazilian translation scholars published an article in order to reflect upon the specificity of Translation Studies in Brazil. In the article, Pagano &

* Professora da Universidade Federal de Campina Grande, Campina Grande, BA

Vasconcellos also mapped TS in the Brazilian context using a CD-ROM published in 2001 which provided a compilation of all research work produced in the Translation Studies (TS) field in Brazil, presenting the theoretical affiliation of each of them between the years 1980 and 1990. The authors used Holmes' (1988) map as a starting point.

Pagano & Vasconcellos (*ibid.*, p. 14) say that, according to Baker (1998, p. 277), this mapping is well-accepted for the organisation of academic activity in the area, since it presents the most traditional lines of research in TS. Its distinction between applied studies (directed to practice) and pure studies (theoretical and descriptive studies developed without any direct practical application) and their subdivisions serve as a guide for translation research. However, for the fact that Holmes' map was developed in the 1970s, there was a need to update his map in order to include some advances from different areas – technology and its tools and research offer, for example. So, the authors (*ibid.*, p. 15) presented a list of sub-categories to be added to Holmes' map.

Translation Studies in Brazil also include 'Applied' and 'Pure' Studies. In the former, the focus of the study is on the computer/translation interface, included in the sub-branch 'Translation Aids' proposed by Holmes; in the latter, this research focuses on Corpus-Based Translation Studies (CTS), included in the sub-branch 'Descriptive'. However, the evolution of technology has provoked the emergence of other types of research, specifically the one being explored here, which becomes a candidate to be added to a specific list – Translation and Online Journalism. The translation of online news reports is an interesting field for research in Translation Studies (TS) as such texts seem to illustrate in an accentuated manner important points at stake in different fields of translation – texts produced in different contexts (American/British ones) reporting issues from specific cultures and realities, in addition to the urgency to translate a text that has to be posted as fast as possible, online.

Relevance of the study of online news reports can be claimed for: (i) the importance of technology in all contexts of life; (ii) the influence of the Internet on the way translations of online news reports are produced, because of the context (generally different cultures and realities), the environment where they are published (the Web); and also (iii) the importance of the computer in the production, dissemination and reception of news events. The computer – and the Internet – is an important tool for text production nowadays and this text production involving technology and the Internet has been problematised in different fields of study – from technology to the human areas.

When we analyse existing target texts, we can never be sure whether a particular translation really reflects a convention or rather the intention of an individual translator. And we would need a rather large number of samples in order to rule out other variables determining the form and quality of a translation, e.g. the (in)competence of the translator or any particular translating instructions given by the initiator. (NORD, 1991b, p. 103)

It seems important to state that the intention here is **not** to focus on any particular linguistic, cultural (or any other specific type of) pattern of translation, but on how the **content** reported is represented from the source texts (ST) selected into the target texts (TT) produced. The choice for content instead of a specific linguistic pattern of translation can be justified still drawing on Nord:

The analysis of text content has not so far been dealt with satisfactorily in the various approaches to translation-relevant text analysis. The concepts of “content”, “meaning”, “sense”, etc. remain vague, and there are very few hints on how to actually elicit the content of a text. The analysis of content is restricted more or less to the level of lexical items (THIEL 1978a, REISS 1984a) and only appears in the form of a summary (THIEL 1978a) or a paraphrase (BÜHLER 1984) of the text. (...)

By “content” we usually mean the reference of the text to objects and phenomena in an extralinguistic reality, which could as easily be a fictitious world as the real world. This reference is expressed mainly by the semantic information contained in the lexical and grammatical structures (e.g. words and phrases, sentence patterns, tense, mood, etc.) used in the text. These structures complement each other, reduce each other’s ambiguity, and together form a coherent context.

Therefore, the starting point for the analysis of content has to be the information carried by the text elements linked on the surface of the text by the text-linguistic linking devices, such as logical connections, theme-rheme relationships, functional sentence perspective, etc. (NORD, 1991, p. 89-90).

Therefore, observing the content of both ST and TT the textual aspects will be examined indirectly via the analysis of the translation strategies used to produce the translations of online news reports.

2 Corpus Selection and Orientation

The aim of this paper is to describe the methodological procedure of a study on the translation of 86 online news reports about Middle East Conflicts observing what Chesterman’s (1997) translation strategies seem to be more frequently used to achieve such a production. When the theme – Middle East Conflicts – and the Corpus – online news reports – were defined, some other important factors for the Corpus selection started to be

considered: (i) the reliability of the sites chosen; (ii) the public interest in relation to the theme; (iii) how such a general theme is reported from one language into another (from American/British English into Brazilian Portuguese, in this case) respecting the possible differences regarding news construction and language structure; and, finally (iv) how the environment (the Web) might influence the translation of online news reports.

As regards online news reports, they offer more dynamic and much faster release of information, demanding, in my opinion, a deeper look and analysis considering aspects related to: news production, audience, context, and the translation practice of online news: who translates the texts, the amount of time set for the translation work to be developed, the space for the specific piece of news available on the site, among others. The point to be raised here involves the interest in translating such pieces of news into different languages – Brazilian Portuguese in this case – and how the act of translating occurs, (i.e., what strategies seem to help translators in their work). In this context, a series of reports about Middle East conflicts are presented in parallel with their translations directed to a specific audience: online journalistic text readers, not necessarily Brazilian readers.

The fact that online news reports are the main interest here it is necessary to investigate some specific criteria adopted for the production of such a text type. Such criteria also influence the production of translation of these texts. Following Nord's (1991) idea that languages play a core role in the process of intercultural communication and translation, I subscribe to the affirmation that languages also help determine and identify who produces the ST; who orders a translation of such texts; the ST in itself; who receives the ST and translates it; the TT in itself; and who reads the TT. The whole process from receiving a ST up to the moment the TT is produced and released is complex, involving several steps and different participants, including readership. The process cannot be related simply to the purpose of the ST, but mainly to how the TT is going to operate in the target context. As Nord (1997, p. 09) says, "the function of the TT is not arrived at automatically from an analysis of the ST, but is pragmatically defined by the purpose of the intercultural communication." Therefore, the hypothesis is that, when translating online news reports, some translations tend to follow a more literal translation type, developed on the linguistic level, while others might show adaptations, additions or omissions, trying to update the TT during the translation process.

3 Corpus-Based Studies and Translation Strategies

Uzar (2002, p. 237) claims that “translation is becoming an increasingly part of society throughout the world”. Because of this, general standard manners of doing translation have improved and, “more and better qualified translations are needed” (*ibid.*). The task of rating how qualified a translation is seems to be related to aspects such as the translator’s expertise, his/her familiarity with the languages involved in the translation process and his/her sense of detachment from the source text in order to perceive among various apparently similar translation possibilities which form fits the target context best. Another helpful aspect is the use of corpora focusing on the investigation of translation products in order to describe specific translation types – in this case, translated online news reports. According to Kenny (2000, p. 93), “the use of real texts and recourse to computational tools to assist in the analysis of these texts are two of the hallmarks of contemporary corpus linguistics”, making the analysis less troublesome.

On the basis of the assumptions and quotes above, this paper focuses on Corpus Linguistics (CL) and Corpus-Based Studies and their application in Translation Studies (TS). The validity of research investigation based on CL combines micro and macro analysis, following what Tymoczko (2002) suggests when she claims that the researcher has two options when approaching research: (i) either looking at the “big picture” – a macroscopic direction –, analysing aspects of culture, for instance; or (ii) analysing particularities of the source and target languages when dealing specifically with translation – a microscopic direction. Tymoczko adds that it is important to take both directions into account when developing a piece of research, because they might be “reinforcing domains to secure strong conclusions” (*ibid.*). What one direction seems to overlook, the other direction will be able to cover more deeply.

The author’s view seems interesting especially for the type of work developed here, because it covers the use of translation tools also taking into account aspects of context and language that might be beyond the scope of such tools. Applying Tymoczko’s view of macro and micro direction to this analysis, what is considered the “big picture” or macroscopic direction in this study corresponds to the three categories Chesterman (1997) creates to classify his translation-strategy framework; and the microscopic direction refers to the strategies themselves:

Chesterman's (1997) three Categories of Translation Strategies

Syntactic strategies	
G1: Literal Translation	“maximally close to the SL form, but nevertheless grammatical.”
G2: Loan, Calque	“A deliberate choice, not the unconscious influence of desired interference.”
G3: Transposition	“any change of word class, from noun to verb, adjective to adverb.”
G4: Unit Shift	A ST unit (morpheme, word, phrase, clause, sentence, paragraph) is translated as a different unit in the TT.
G5: Phrase Structure Change	“a number of changes at the level of the phrase, including number, definiteness and modification in the noun phrase, and person, tense and mood in the verb tense.”
G6: Clause Structure Change	“changes that have to do with the structure of the clause in terms of its constituent phrases.”
G7: Sentence Structure Change	“it affects the structure of the sentence unit.”
G8: Cohesion Change	“it affects intra-textual reference, ellipsis, substitution, pronominalization and repetition, or the use of connectors of various kinds.”
G9: Level Shift	“the mode of expression of a particular item is shifted from one level (phonology, morphology, syntax and lexis) to another.”
G10: Scheme Change	“kinds of changes that translators incorporate in the translation of rhetorical schemes such as parallelism, repetition, alliteration, metrical rhythm, etc.”
Semantic Strategies	
S1: Synonym	“selects not the obvious equivalent but a

	synonym or near-synonym for it.”
S2: Antonym	“the translator selects an antonym and combines this with a negation element.”
S3: Hyponymy	“shifts within the hyponymy relation are common.”
S4: Converses	“pair of (usually) verbal structures which express the same state of affairs from opposing viewpoints, such as buy and sell.”
S5: Abstraction change	“a different selection of abstraction level may either move from abstract to more concrete or from concrete to more abstract.”
S6: Distribution change	“change in the distribution of the ‘same’ semantic components over more items (expansion) or fewer items (compression).”
S7: Emphasis change	it “adds to, reduces or alters the emphasis or thematic focus, for one reason or another.”
S8: Paraphrase	It “results in a TT version that can be described as loose, in some contexts even undertranslated. Semantic components at the lexeme level tend to be disregarded, in favour of the pragmatic sense of some higher unit such as a whole clause.”
S9: Trope change	It “applies to the translation of rhetorical tropes (i.e. figurative expressions).”
S10: Other semantic changes	“include other modulations of various kinds, such as change of (physical) sense or of deictic direction.”
Pragmatic strategies	
Pr1: Cultural filtering	“also referred as naturalization, domestication or adaptation.”
Pr2: Explicitness change	“either towards more explicitness (explicitation)

	or more implicitness (implication).”
Pr3: Information change	“either the addition of new (non-inferable) information which is deemed to be relevant to the TT but not present in the ST, or the omission of ST information deemed to be irrelevant.”
Pr4: Interpersonal change	“it alters the formality level, the degree of emotiveness and involvement, the level of technical lexis and the like: anything that involves a change in the relationship between text/author and reader.”
Pr5: Illocutionary change	“linked with other strategies: changing the mood of the verb from indicative to imperative, a change from statement to request.”
Pr6: Coherence change	“the logical arrangement of information in the text, at the ideational level.”
Pr7: Partial translation	“any kind of partial translation, such as summary translation, transcription, translation of the sounds only, and the like.”
Pr8: Visibility change	“a change in the status of the authorial presence, or to the overt intrusion or foregrounding of the translatorial presence. For instance, translator’s footnotes, bracketed comments or added glosses explicitly.”
Pr9: Transediting	“the sometimes radical re-editing that translators have to do on badly written original texts.”
Pr10: Other pragmatic changes	Changes in the layout of the text, for example; or in the dialect choice.

Chesterman’s first category – Syntactic Strategies – deals specifically with the structure and organisation of the phrases; the second category – Semantic Strategies –

concentrates on the meaning of sentences associated with the lexical choice, based on syntactic terms that best fit in a certain context; and the last category – Pragmatic Strategies – goes beyond syntactic and semantic aspects and deals with meaning in terms of what content is important to be present in the TT.

3.1 The Focus on Memes in place of Norms

Chesterman (1997, p. 02) uses the term ‘memes’ as a metaphor for translation explaining that a meme “is simply an idea that spreads”. The author (*ibid.*, p. 51) adds that “[i]f a meme comes to dominate (for any reason: practical, political, cultural, aesthetic...), and competing memes fade, one course of development is that such a meme becomes regarded as a norm – whether imposed by an authority or simply accepted as such.”

Chesterman (*ibid.*) makes a distinction between the terms ‘Normative’ and ‘Prescriptive’ saying that genuinely scientific translation theory (Translation Studies – TS) is descriptive, and, consequently, normative. Applied research – or translation training – on the other hand, focuses on what translations should be like, prescriptively. Here, the focus is on descriptive work, not on training or prescribing translation procedures. The author uses the terms ‘norm’ and ‘normative’ in a descriptive sense, in order to focus on the description, understanding and explanation of translation norms. The author adds that a certain community inevitably shares common ideas regarding what is ‘correct’ of a certain act or behaviour and the group agrees on what is in some sense ‘right’ or ‘wrong’.

Chesterman (1997, p. 55-56) presents the function of norms as being evolutionary. The author says that norms regulate behaviour, being beneficial to all parties, creating and maintaining social order, facilitating material and social interaction, and even facilitating cognitive processing. The author goes on to say that norms save both time and effort, of great importance in the world today, given the fact that society lives urgent times, demanding fast solutions, and fast decisions.

4 Using Corpora in Translation Studies Methodologically

Kenny (2000) states that the first thing to be done when developing a corpus-based investigation is to specify exactly what the researcher wants to study. After determining the ‘what’ very clearly, it is possible to start thinking about how the investigation will be

undertaken. Kenny (*ibid.*, p. 94) says that “it is clear that how we ask the research question has implications for the kind of corpus resources needed in any particular study.” The research questions thus determine whether the researcher will opt for a parallel corpus analysis or any other type. Here, the option is the parallel corpus analysis, since the work is carried out to identify translation strategies used in translations of online news reports. The parallel-corpus approach will help to find the strategies in the target texts (TT) that are of particular interest.

Kenny (*ibid.*) goes on to say that “the first challenge for the researcher is thus the specification of such interesting points in the source texts (ST).” I subscribe to this comment and add that, in fact, another challenge is exactly what to select from the corpus, because in some cases several different aspects of a specific text type seem to be interesting and worth being investigated. The parallel-corpus approach, to my view, helps the researcher focus on what was asked in the research questions without feeling lost among so much information and so many interesting points of investigation. The next step according to the author is then parsing and tagging the corpus in order to start marking what is being looked for – here, the translation strategies used in the TTs. When I first started aligning and tagging the TT and ST in order to identify Chesterman’s translation strategies, some points immediately started to call my attention: (i) sometimes the length of the TTs was smaller than the STs, or vice-versa; (ii) sometimes the TT started with what is considered to be a typical translation, with similar ideas in sentences and paragraphs, containing synonyms and similar semantic ideas. As I continued reading, I noticed that a different text started to be produced inside the translation. In other words, a TT suddenly started to present ideas, sentences and paragraphs that weren’t presented in the ST. Such a fact made me decide to include more sites to compare with the translations appearing in NYT/UOL. That was when I included BBC News and Reuters, which have Brazilian Portuguese versions of their sites, making me select the TT and ST at the same time.

It was noticed that both the ST and the TT on *BBC News* and Reuters are updated in terms of content several times during the day, meaning that there is apparently no final version for the reports. This point reminds us of the importance of marking not only the date but mainly the time of publication of the report, besides including the observation concerning the addition of information. In the case of *BBC News* and *Reuters*, the headlines were much closer in meaning to the headlines of the ST, which facilitated my search. On the other hand,

the reports presented more modifications than the ones on NYT/UOL. By modifications I mean: the addition or omission of information, which made the texts much longer or much shorter than the ST. Sometimes, on the three sites, one sentence in the TT was a translation from a sentence in the ST, but the location of the sentences was not aligned because of the inclusion of information, making the identification more painstaking. The translation strategies were identified in order to analyse the corpus as well as to observe if the translations produced followed a more target-text oriented aspect, giving emphasis to necessary adaptations and to the target readership, or if they were produced following a more source-text oriented practice of translation, close to syntactic and semantic features of the source text.

4.1 Type of Corpus

Zanettin (2000) describes different types of ‘translation-driven’ corpora and elaborates on their design and analysis. The author mentions the constant interest in developing research using computer-assisted methods of investigation in TS and moves on to cite the types of ‘translation-driven corpora’:

- “the monolingual comparable corpus, consisting of a set of translations and a comparable set of texts spontaneously created in the same language and selected according to similar criteria” (p. 105);
- “the bilingual comparable corpus. Neither of the two components includes translated texts; what is compared are texts spontaneously produced in two languages under similar circumstances and within the same domain” (p. 106). For this second type, the author explains that its purpose is “to develop a tool and a resource for trainees and practitioners in the translation profession” (*ibid.*). – Obviously, this is not the case under study here, hence it was not opted for.
- The parallel corpus: “a set of translations in one language and their respective source texts in another language” (*ibid.*).

4.2 Corpus Representativeness

Baker (1995) describes the following criteria for corpus design:

- Machine-Readable: collection of texts held in electronic form, read and analysed automatically or semi-automatically;
- Full-Text: full texts and not text fragments allowing the examination of micro level units such as words, phrases and sentences;
- Open-endedness: corpus designed with a view to the future, as an open-ended entity in which texts are constantly added to; and
- Representativeness: in building the corpus in a specific area, the researcher knows the extent to which and in what respects this corpus is representative enough to serve its purpose. The author adds that selecting a representative corpus is related not only to size, but also to the careful description of what the corpus is intended to represent.

I believe the corpus selected fits in the description presented by Baker, and I subscribe to the author's last comment because my interest is much more directed to qualitative aspects concerning the corpus than to quantitative aspects, which will serve as support for the elaboration of a qualitative analysis.

Having said that, I present what Zanettin states about the point under discussion:

One of the major issues in corpus design is that of representativeness; what distinguishes a corpus from a collection of electronic texts (or a text archive) is that a corpus is put together in a principled way so as to be representative of a larger textual population, in order to make it possible to generalise findings concerning that population. Thus, the most appropriate design for a corpus depends on what it is meant to represent. (*apud.* BIBER 1993; HALVERSON 1998; KENNEDY 1998; BIBER et al. 1998). This should be remembered before making any general statements about language, texts, or translations based on corpus analysis; what is found in a corpus will only apply to what that particular corpus represents. (ZANETTIN, 2000, p. 107).

Zanettin's quote and Baker's description seem to have points in common, and will thus serve as the basis for the Corpus selected here. Moreover, elaborating on the issue of representativeness, Halverson (1998) cites Biber's discussion on data collection:

Typically researchers focus on sample size as the most important consideration in achieving representativeness: how many texts must be included in the corpus, and how many words per text sample. Books on sampling theory, however, emphasise that sample size is not the most important consideration in selecting a representative sample; rather, a thorough definition of the target population and decisions concerning the method of sampling are prior considerations. (BIBER, 1993, p. 243).

For the selection of texts to build the Corpus under study, I have focused on both directions: the Corpus size – 86 texts, being those 43 target texts and their respective source texts aligned in parallel form – as well as the research questions to be answered:

- 1) What translation strategies are most frequently used when translating online news reports?
- 2) What are the implications of the use of the translation strategies identified in the TT?
- 3) Has the environment in which the source and target texts of the corpus were published (i.e. the Web) influenced the translation results and if so, how?

4.3 The Selection of Texts: The Target Population

Halverson (1998, p. 03) states that “defining a target population is paramount” when defining the corpus and selecting its texts. The author complements this point with Biber’s (1993) statement about the two aspects of the definition of a target population: 1) the boundaries of the population – what texts are included or excluded from the population; 2) hierarchical organisation within the population – what text categories are included in the population, and what are their definitions.” Halverson (*ibid.*, p. 03) adds that a target population “consists of specifying the boundaries and internal structure of the category” [opted for the corpus construction]. Such boundaries will help determine which instances must be included and which ones must be left aside.

4.4 Corpus Description

Zanettin (2000, p.115) affirms that “parallel corpora allow not only quantitative analysis (...). They also facilitate qualitative analysis (...).” The author mentions the need to align the Corpus in order to carry out the investigation, adding that “alignment procedures can, to a large extent, be automated, and may be performed on the basis of statistical elaboration, taking into account the number of sentences, words or even characters in the pairs of texts to be aligned.” However, the author is aware of the fact that there must be a degree of interaction between machines and humans in order to “provide a way of examining in some detail how translations map onto source texts” (*ibid.*).

5 Corpus Analysis: The Use of Translation Strategies in the Parallel Corpus

The objective here is to investigate the use of translation strategies in the Corpus selected and to show the results of the investigation presenting some examples. According to the general results, Strategy G1 (literal translation) was highly used, followed by S1 (synonym), Pr3 (information change), Pr4 (interpersonal change) and Pr2 (explicitness change), respectively, as demonstrated in the next examples:

Example 01: (TT): [Muitas vezes já se disse que a verdade é a primeira baixa da guerra G1, G5]. [Durante uma campanha eleitoral para a presidência, isso pode ser mais verdadeiro que nunca G1 S1]. [Considere uma pergunta aparentemente simples: Qual é o custo da guerra do Iraque para os Estados Unidos G1]? [O presidente Bush e o senador John Kerry deram respostas diferentes, mas ambos os candidatos ignoraram o que talvez seja o item mais caro: o impacto da guerra sobre a economia em geral G1].

(ST): [It's often said that truth is the first casualty of war.] [During a presidential campaign, that may be more apt than ever.] [Consider a seemingly simple question: What is the cost of the Iraq war to the United States]? [President Bush and Senator John Kerry have given different answers, but both candidates have ignored what may be the biggest cost item: the war's impact on the overall economy]. (NYT/UOL – text 1)

Example 02: (TT): [O ataque ocorreu na mesquita Shahidain, que é cercada de moradias baratas na região de Tameem, no centro de Mosul G1 S1 Pr2]. [Testemunhas disseram ter ouvido um forte estrondo dentro do pátio da mesquita, que ainda está em construção G1 Pr2 Pr3 Pr8].

(ST): [The attack was at the Shahidain mosque which is surrounded by cheap housing in Mosul's central Tameem neighbourhood.] [Witnesses described seeing a ball of fire and hearing a huge explosion inside the courtyard of the mosque, which is still under construction. (BBC – text 1)]

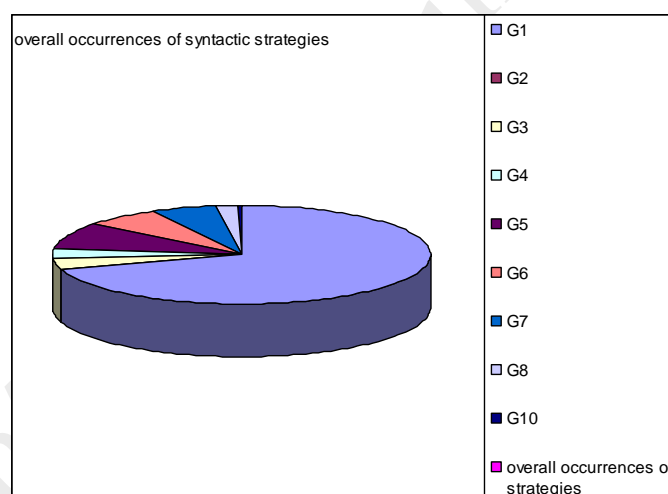
Example 03: (TT): [Kerry e muitos dos outros críticos do presidente argumentam que sua adoção de uma democratização liderada pelos Estados Unidos - repleta de uma calorosa referência a Harry Truman, o presidente que iniciou a reconstrução da Europa e do Japão - representa pouco mais do que uma justificativa pós-fato para a guerra G1 S1 Pr4]. [Eles notam que Bush fez apenas um grande discurso sobre a democratização do Oriente Médio antes de invadir o Iraque -apesar de ter falado quase diariamente sobre a ameaça das armas de destruição em massa Pr4 G1 S1]. [("Nós precisávamos de mais alguns discursos sobre a democratização, e menos do outro", reconheceu posteriormente um de seus principais conselheiros, no ano passado) Pr3 G6]. [Agora, faz parte de sua mensagem diária G1]. [Os críticos argumentam que o discurso de Bush encobre todos os erros dos últimos 18 meses, que dificultaram para os reformadores na região semear as sementes da mudança Pr4 G1].

(ST): [Mr. Kerry and many of the president's other critics argue that his embrace of American-led democratization - replete with a warm reference to Harry S. Truman, the president who initiated the reconstruction of Europe and Japan - amounts to little more than an ex post facto justification of the war]. [They note that Mr. Bush gave only one major speech about democratizing the Middle East before invading Iraq, though he spoke almost daily of the threat of unconventional weapons. [("We needed a few more of the democratization speeches, and less of the other," one of his most senior advisers conceded late last year.)] [Now, it is part of his daily message]. [Critics argue that Mr. Bush's speech glosses over all the mistakes of the last 18 months that have made it more difficult for reformers in the region to sow the seeds of change]. (The NYT – text 3)

There is a linguistic progression linked to this distribution, for example, the Syntactic Strategies evolve from word level (G1 and G2), to phrase/clause/sentence structure change (G3 to G7), and then, they evolve to cohesion and level of language such as morphology and phonology. The same happens to the Semantic Strategies, also evolving from word level (S1 to S4) to changes in the semantic level of abstraction, distribution of semantic components and emphasis (S5 to S7). After that, there is the use of paraphrase (G8). The Syntactic and Semantic categories seem to overlap with each other, and this overlapping can be observed in

the organisation and distribution presented. The Pragmatic Strategies do not follow this pattern, because they do not have a direct relation to the linguistic system of languages as the Syntactic and Semantic Strategies do. The Pragmatic Category is based on the analysis of a specific agenda, i.e., specific interests concerning clients, context, translators, editors, time constraints etc., regarding the application of its strategies and also on what was not possible to fit in the Syntactic or Semantic Category, marking the translator's interference or some kind of deviation. Strategies G10, S10, and Pr10 try to be broad enough to cover any gap not possible to be answered by the other strategies.

There is an extreme discrepancy regarding the frequency of use among the strategies. Category 1 – Syntactic Strategy – was predominant in the Corpus in terms of frequency of use. However, this predominance occurs because of the constant use of one single strategy – G1 – in the texts. This predominance of G1 is so repetitive that it places the Syntactic category as the most used one in the whole Corpus. Nevertheless, some Syntactic Strategies are hardly used, for example, G2 and G10, as can be observed next:



Pie 1. Overall occurrences of Syntactic Strategies in the Corpus.

Pie 1 demonstrates the frequent use of Syntactic Strategies, generating more literal translations. They seem to be, in fact, a very common option for translators of online news reports.

5.1 Strategy Overlapping

As the strategies evolve, they need to be complemented, in some cases by strategies from the same or any other of the two categories. The most common overlapping of strategies

happened between G1/S1; G1/S6; G1/G5 and Pr3/Pr8. It is possible to explain these combinations for the fact that, even using a syntactic, literal translation strategy – G1 –, it is necessary to make adjustments from ST to TT due to linguistic differences, as demonstrated in the next examples:

Example 04: TT: ["Mas tal progresso está ocorrendo mais rapidamente do que muitos diriam ser possível G1 S1]. [Por toda uma região conturbada, nós estamos vendo um movimento na direção de eleições, maiores direitos para as mulheres e uma discussão aberta sobre reformas pacíficas G1 S1]. [A eleição no Afeganistão há menos de duas semanas foi um evento marcante na história da liberdade G1]. [Tal eleição foi uma derrota tremenda para os terroristas G1 S1"].

ST: ["Yet, that progress is coming faster than many would have said possible]. [Across a troubled region, we are seeing a movement toward elections, greater rights for women, and open discussion of peaceful reform]. The election in Afghanistan less than two weeks ago was a landmark event in the history of liberty]. [That election was a tremendous defeat for the terrorists."] (NYT/UOL – text 3)

Example 05: TT: [**Rohani afirmou que o Irã vai "retomar parte das atividades nucleares em um futuro próximo"**, mas que **"as condições e a hora" da iniciativa ainda estavam sendo discutidas com os negociadores da França, da Alemanha e da Grã-Bretanha. G1 S6]**

ST: ["If Iran cannot exercise its rights with in the framework of the NPT, it will no longer have any respect for this treaty," Mr Rohani said.] [**"Iran will definitely resume a part of its enrichment activities in the near future... but we are still discussing its conditions and time of restarting the activities."**] (BBC – text 25)

Example 06: TT: [Mas, num sinal de tensões à frente, forças israelenses **fizeram buscas na cidade de Jenin, na Cisjordânia**, na segunda-feira e prenderam oito militantes palestinos, **acusados de preparar ataques G1 S6.**]

ST: [But in a sign of further tension, Israeli forces **raided the West Bank city of Jenin** Monday and arrested eight Palestinian militants they said **were making rockets for future attacks.**] (Reuters – text 3)

Example 07: TT: [Aquele contrato de um ano não foi renovado, mas a empresa **já estava conseguindo** outros, diretamente junto à Autoridade Provisória da Coalizão ou como empresa subcontratada de outras G1 G5].

ST: That one-year contract was not renewed, but the company **had already begun pulling** in others, directly with the Coalition Provisional Authority or as a subcontractor to other companies]. (NYT/UOL – text 2)

Example 08: TT: [Mas **ele afirmou** que **não vai negociar** que tecnologia o país usará em seu programa nuclear. G1, G5]

ST: "The Iranian nation **will never hold negotiations** about its definite rights," **he said**. (BBC – text 37)

Example 09: TT: ["Demos uma ordem para planejar a transferência de Qalqilya, mas não para transferi-la", disse G1 G5.]

ST: ["We gave an order to plan the way Qalqilya will be transferred, but not to transfer it yet," he said.] (Reuters – text 1)

Example 10: TT: [O líder curdo Jalal Talabani será nomeado o novo presidente do Iraque em uma reunião do Parlamento na quarta-feira, afirmaram fontes do governo. Pr3 Pr8]

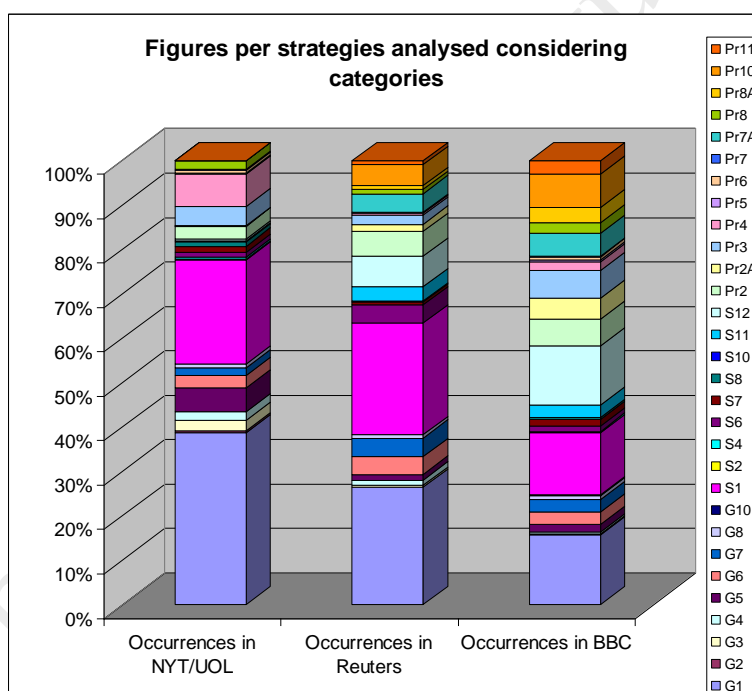
ST: [Leaders of the main political blocs said veteran Kurdish politician Jalal Talabani would be named president at a parliament meeting on Wednesday, more than nine weeks after millions of Iraqis braved insurgent bombs to vote.] (BBC – text 4)

Example 11: [Três funcionários do alto-escalão do Departamento da Defesa explicaram nesta quarta-feira (9) que as novas instruções evidenciam a proibição da utilização de cães com focinheiras durante os interrogatórios G7 Pr8]. [Segundo eles, as instruções fornecem orientações específicas para as unidades que atuam no terreno, tais como o período de tempo máximo durante o qual elas podem manter os prisioneiros sob custódia antes de libertá-los ou encaminhá-los para os seus superiores hierárquicos, os quais, então, determinarão as condições de sua detenção Pr3 Pr8].

ST: [Three senior defense officials said Wednesday that the new procedures clarified the prohibition against the use of muzzled dogs in interrogations, gave specific guidance to field units as to how long they could hold prisoners before releasing them or sending them to higher headquarters for detention, and made clear command responsibilities for detainee operations]. (The NYT – text 4)

It seems important to reinforce the idea that on *UOL* the TTs seem to follow a more source-text oriented pattern, meaning that the translations present a static pattern, without frequent changes or adaptations from the ST to the TT, while on the other sites, there was a tendency to have TTs frequently transformed into a new online journalistic text, showing a more dynamic pattern.

The analysis carried out here has provided the following figures:



Graphic 1. Specific Occurrences in NYT/Reuters/BBC.

The translators from *UOL* seem to have favoured the use of syntactic strategies, and have opted for, in order of frequency, G1 (literal translation), S1 (synonym), Pr4 (interpersonal change), G5 (phrase structure change), Pr3 (information change), Pr2 (explicitness change), G6 (clause structure change), and G3 (transposition). The tendency here was to recur to the use of syntactic strategies, opting for G5 (phrase structure change) and G6 (clause structure change).

The *BBC* site presented translated texts that were more dynamic, having part of the paragraphs translated from the ST, part omitted and several pieces of information added to it. In some cases, these TTs from the *BBC* site were presented as a kind of summary of the whole ST piece of news. Some translation strategies were not used, namely: G2 (loan, calque), S2 (antonym), and S10 (other semantic changes). Strategies G1 (literal translation), S1 (synonym), Pr 10 (other pragmatic changes), Pr3 (information change), and Pr2 (explicitness change) were, respectively, the most commonly used.

This panorama shows that the *BBC* site differs from the *NYT* site in terms of focusing more on the use of semantic and pragmatic strategies as well as presenting more dynamic texts. Here, the predominance of pragmatic strategies was higher than on the *NYT*, which shows that the translators/journalists indeed produced, consciously or unconsciously, more target-text oriented translations. *Reuters* was the site that presented more changes in terms of omissions and additions from the ST to the TT. This fact, just like on the *BBC* site, shows a greater tendency to follow the *skopos* of the supermeme called ‘most-up-to-date information’. This supermeme focuses on the updating information from the ST to the TT from time to time. Strategies G2 (loan, calque), G10 (scheme change), S2 (antonym), S4 (converses), S10 (other semantic changes), Pr6 (coherence change) and Pr7 (partial translation) were not used. The strategies most commonly used were G1 (literal translation), S1 (synonym), Pr2 (explicitness change), Pr10 (other pragmatic changes); G7 (sentence structure change), S6 (distribution change), and G6 (clause structure change), respectively. *BBC* and *Reuters* therefore were the sites that presented translations that could be described as dynamic translations, because there was a higher number of changes on the focus of information provided by the use of the strategies Pr2 (Explicitness change), Pr3 (Information change) and Pr7 (Partial translation), besides the fact that the translator/journalist was in a position that favoured him/her to make changes from ST to TT apparently with more freedom, marking the use of the strategy Pr8 (Visibility change).

As pinpointed previously, strategy G5 (phrase structure change), for example, was generally opted for on *UOL Jornais* and rarely used on *BBC News* or *Reuters*, because, in order to produce the sentences and ideas in the translations at *UOL Jornais*, it was necessary to use more nominalisations and verb changes, while on *BBC News* and *Reuters* the sentences were, in fact, retextualised, and the following strategies prevailed strategies G1 (literal

translation), S1 (synonym), Pr2 (explicitness change), Pr3 (information change), and Pr10 (other pragmatic changes), as observed in the next example:

Example 12: (TT): [Pelo menos 14 pessoas morreram na explosão de um carro-bomba num mercado lotado em Suwayrah, ao sul da capital iraquiana, Bagdá. G5 S1 Pr2 Pr3]
[Segundo as autoridades, 43 pessoas ficaram feridas e o número de vítimas fatais deve subir. Pr2 Pr3]

[A cidade é um conhecido reduto de insurgentes numa região apelidada de Triângulo da Morte do Iraque. Pr2]

[Autoridades iraquianas anunciaram nesta sexta-feira que os corpos de 14 homens foram encontrados em uma vala comum em Bagdá. G5 Pr2 Pr3 Pr10]

(ST): [At least 16 people have been killed in a car bomb attack in a market south of Baghdad, Iraqi police say.]

[As many as 40 people were injured in the blast at Suwayra, about 60km (38 miles) south of the capital.]

In Baghdad, police found a shallow grave with the bodies of 14 men, apparently the victims of execution-style killings.

[Meanwhile at least eight police officers died in a bombing near the northern city of Tikrit.] There has been intense violence since a new Iraqi government was announced in late April. More than 250 people have died.

In the Suwayra attack, a car exploded near a crowded vegetable market at about 1500 (1100 GMT) on the Friday holiday.

Mixed town

Police said the blast blew away stalls and destroyed cars.

Doctors at the local hospital said the most severely wounded had been transferred to hospitals in Baghdad.

Suwayra is a mixed Shia and Sunni Muslim town, but is near the so-called Sunni triangle, where there have been many attacks on Iraqi and coalition forces.

[Earlier in Baghdad, police investigated a hole in the ground and found 14 bodies, dressed in the long white robes favoured by Sunni Islamists.] (BBC – text 20)

Reuters and *BBC* differ from *NYT/UOL* in that they use strategies that adapt the TT because of the need to update information constantly on the sites. These sites follow the specific purpose of reporting what happened or what was said some hours/instants ago.

5.2 Dynamic/Static Translations

The figures on *UOL* show that syntactic strategies were most frequently used. Nevertheless, semantic and pragmatic strategies were used as well. The implication of a predominance of syntactic strategies is exactly that the translated texts produced tended to be static, i.e. they were not updated after being posted on the site, and followed a fixed source-text pattern.

Moving to the *BBC* and *Reuters* sites, they tended to use less syntactic strategies in comparison to the *UOL* site, opting for more semantic and pragmatic strategies. The texts on these two sites were considered dynamic texts, for the fact that they were constantly updated after being posted on the sites and followed a target-text pattern. It seems important to highlight that strategies G9 (level shift), S3 (hyponymy), S5 (abstraction change), S9 (trope change), Pr1 (cultural filtering) and Pr9 (transediting) were not used in the analysed translations. These strategies are more likely to be used in literary texts, for example, or texts having cultural aspects concerning the language and life of specific peoples, which was not the case here either. In informative texts, attention is drawn upon facts. When cultural terms come up, they are described or explained in a more pragmatic form.

6 Final Remarks

Researchers in the area of TS have worked hard to store examples from varied corpora in order to illustrate established and described theories and also to look for answers or suggestions on how to produce a specific translation type with higher assurance. The use of corpora in TS has demonstrated to be valid and necessary in order to provide high-quality research in terms of representativeness and results. Another positive point about the use of corpora in TS is that it is rewarding to come to a moment in which the researcher starts to see the results after the compilation, organisation and investigation of the corpus.

The use of the term ‘meme’ denotes the idea that decisions are made after a process of ‘selection of translation options’ and the surviving memes, i.e. the most commonly used

memes, grow strong, while the rest of them dies out. In fact, translators should follow a procedure based on concrete data, making choices based on sensible behaviour. These points are basically directed to the translation of the texts.

The motivation underlying the translation tendencies emerging from the data analysis can be related to basically three factors: (i) the *skopos* opted for and determined by the news agency; (ii) time constraints that influence the publishing of online news as fast as possible; and (iii) the need to keep readers informed by as many pieces of news as possible, almost at the same time. What I noticed after analysing the results emerging from the data analysis was that the Web and the process of Online Journalism was the main influence for the translation tendencies found here: static memes (source-oriented translations with no updates after posted on the site) and dynamic memes (target-oriented translations which are constantly updated). These tendencies should not be considered as right or wrong. Both of them serve different *skopoi* and are able to achieve the aims determined for them. The only consideration I would make is that dynamic memes serve the objective of producing updated news, which is the *skopos* that goes beyond the production of a static translation that will suffer no further changes. Therefore, journalists must take advantage of new technological advancements in order to know how to use them in their favour. I would like to add that the same should be said concerning translators. In fact, when translating online news reports, the role of a journalist and a translator almost inevitably will overlap.

7 Bibliographical References

BAKER, M. Corpora in Translation Studies: An Overview and some Suggestions for Future Research. In: **Target**, 7. Amsterdam: John Benjamins, 1995, p. 223-243.

BIBER, D. Representativeness in Corpus Design. In: **Literary and Linguistic Computing**, 8. Australia: Cambridge University Press, 1993, p. 243-257.

CHESTERMAN, A. **Memos of Translation**: The Spread of Ideas in Translation Theory. Amsterdam: John Benjamins, 1997, p. 87-116.

DEUZE, M. Understanding the impact of the Internet: On new media professionalism, mindsets and buzzwords. In: *EJournalist*, volume 1, number 1, at <http://www.ejournalism.au.com/ejournalist/deuze.pdf>. 2001.

HALVERSON, S. Translation Studies and Representative Corpora: Establishing Links between Translation Corpora, Theoretical/Descriptive Categories and a Conception of the Object of Study. In: **Meta**, XLIII, 4. Montreal: Erudit, 1998, p. 01-22.

KENNY, D. Lexical Hide-and-Seek: Looking for Creativity in a Parallel Corpus. In: Olohan, M. **Intercultural Faultlines: Research Models in Translation Studies I: Textual and Cognitive Aspects**. UK: St. Jerome, 2000, p. 93-104.

NORD, C. **Translating as a Purposeful Activity: Functionalist Approaches Explained**. Manchester: St. Jerome, 1997.

TYMOCZKO, M. Computerized Corpora and the Future of Translation Studies. In: **Meta, XLIII, 4**. Montreal: Erudit, 1998, p. 1-15.

UZAR, R. S. A Corpus Methodology for Analysing Translation. **Cadernos de Tradução, IX**. Florianópolis: UFSC, 2002, p. 237-265.

ZANETTIN, F. Parallel Corpora in Translation Studies: Issues in Corpus Design and Analysis. In: OLOHAN, M. **Intercultural Faultlines: Research Models in Translation Studies I: Textual and Cognitive Aspects**. UK: St. Jerome, 2000, p. 105-118.

<http://www.nytimes.com>

<http://news.bbc.co.uk>

<http://noticias.uol.com.br/jornais>

<http://www.reuters.com>