Meta
Journal des traducteurs
Translators' Journal

META

A Corpus-Based Evaluation Approach to Translation Improvement

Ghodrat Hassani

Volume 56, numéro 2, juin 2011

Les corpus et la recherche en terminologie et en traductologie
Corpora and Research in Terminology and Translation Studies

URI : https://id.erudit.org/iderudit/1006181ar
DOI : https://doi.org/10.7202/1006181ar

Aller au sommaire du numéro

Éditeur(s)

Les Presses de l'Université de Montréal

ISSN

0026-0452 (imprimé)
1492-1421 (numérique)

Découvrir la revue

Citer cet article

Hassani, G. (2011). A Corpus-Based Evaluation Approach to Translation Improvement. *Meta*, *56*(2), 351–373. https://doi.org/10.7202/1006181ar

Résumé de l'article

En contexte professionnel, l'évaluation des traductions s'est toujours heurtée au problème de la subjectivité, au détriment des évaluateurs, qu'ils soient clients, traducteurs ou fournisseurs de service. Il semble qu'il devienne possible d'alléger ce fardeau, de telle manière que les évaluateurs gagnent en objectivité et les traducteurs retirent un bénéfice à améliorer leurs compétences – et leur statut professionnel – à la lumière des commentaires que leur font constamment ceux qui évaluent leurs traductions. L'objectif du présent travail est d'explorer les voies prometteuses qu'ouvre l'évaluation fondée sur des corpus. Faisant appel à un corpus d'anglais américain contemporain (COCA) à des fins d'évaluation en contexte professionnel, la stratégie mise en oeuvre dans la présente étude envisage l'évaluation des traductions comme un moyen permettant d'atteindre un objectif pertinent, c'est-à-dire l'amélioration de la qualité de la traduction. Elle illustre également comment les caractéristiques spécifiques d'un corpus peuvent atténuer les composantes subjectives de l'évaluation, ce qui permet d'obtenir, *in fine*, des traductions de qualité supérieure.

A Corpus-Based Evaluation Approach to Translation Improvement

GHODRAT HASSANI
Allameh Tabataba'i University, Tehran, Iran
qhassani@gmail.com

RÉSUMÉ

En contexte professionnel, l'évaluation des traductions s'est toujours heurtée au problème de la subjectivité, au détriment des évaluateurs, qu'ils soient clients, traducteurs ou fournisseurs de service. Il semble qu'il devienne possible d'alléger ce fardeau, de telle manière que les évaluateurs gagnent en objectivité et les traducteurs retirent un bénéfice à améliorer leurs compétences – et leur statut professionnel – à la lumière des commentaires que leur font constamment ceux qui évaluent leurs traductions. L'objectif du présent travail est d'explorer les voies prometteuses qu'ouvre l'évaluation fondée sur des corpus. Faisant appel à un corpus d'anglais américain contemporain (COCA) à des fins d'évaluation en contexte professionnel, la stratégie mise en oeuvre dans la présente étude envisage l'évaluation des traductions comme un moyen permettant d'atteindre un objectif pertinent, c'est-à-dire l'amélioration de la qualité de la traduction. Elle illustre également comment les caractéristiques spécifiques d'un corpus peuvent atténuer les composantes subjectives de l'évaluation, ce qui permet d'obtenir, *in fine*, des traductions de qualité supérieure.

ABSTRACT

In professional settings translation evaluation has always been weighed down by the albatross of subjectivity to the detriment of both evaluators as clients and translators as service providers. But perhaps this burden can be lightened, through ongoing evaluator feedback and exchange that foster objectivity among the evaluators while sharpening the professional skills and recognition of the translators. The purpose of this paper is to explore the promising avenues that a corpus-based evaluation approach can possibly offer them. Using the Corpus of Contemporary American English (COCA) for evaluation purposes in a professional setting, the approach adopted for this study regards translation evaluation as a means to a worthwhile end, in a nutshell, better translations. This approach also illustrates how the unique features of the corpus can minimize subjectivity in translation evaluation; this in turn leads to translations of superior quality.

MOTS-CLÉS/KEYWORDS

évaluation de la traduction, milieux professionnels, erreurs de traduction, approche fondée sur les corpus, amélioration de la traduction
translation evaluation, professional setting, translation errors, corpus-based approach, translation improvement

Introduction

Despite a gallimaufry of material with respect to the concept of translation quality assessment, it has, to use Maier's words (2000: 137), "historically been particularly vexed." The highly subjective and arbitrary nature of translation evaluation, the elusiveness and slipperiness of a conclusive answer to the question of what a good or

bad translation is, and, according to Farahzad (2003: 30), the fine line that translator trainers have to walk between *contrast* and *criticism* have all aggravated this vexation.

Regarding the question of subjectivity in translation quality assessment, despite the strenuous efforts of some scholars (Reiss 1971/2000; House 1981) to lay down firm and standard criteria for evaluation, the nagging pain of subjectivity still hangs tough like the sword of Damocles and appears to be here to stay. Therefore, carrying out further research in the hope of establishing a set of objective standards for rigorous and thoughtful analysis sounds sensible. Also, addressing the question of "how to tell whether a translation is good or bad," House (1993: 4700) points out that "any statement that describes the positive and negative features of a given translation, arriving at a summative assessment of its worth, implies a conception of the nature and goals of translation." Particularly significant in the case of translation quality assessment is the evaluator's own stance on translation itself. For a case in point, if the evaluator is a literalist, she/he would be on the lookout for the rubrics and guidelines as laid out for a literalist translator. If the given translation lives up to those rubrics and guidelines, it would be a good translation, otherwise, inadequate at best. Finally, Farahzad (2003: 30) tackles the fuzzy boundaries between *contrast* and *criticism*, saying that *contrast* is sometimes mistaken for *criticism* so that the target text is compared with the source text *piece by piece* and the translator's mistakes are accounted for in regard to semantics, grammar, equivalence, and so on, without access to any theoretical framework of translation criticism. She regards such an approach to translation criticism as highly *reductionist* and believes that this approach falls badly short of providing a solid basis for the evaluation or criticism of a given translation. She finds the good of this approach in teaching translation and, like House (1998: 197), takes for granted a theoretical framework for the purpose of evaluating translations.

The urgency of remedying these fundamental shortcomings has galvanized scholars like Bowker (2000; 2001) into adopting a corpus-based approach to evaluating translations so that the subjective element could be reduced to a minimum as much as possible and evaluators could have access to a cornucopia of authentic texts as a benchmark against which the goodness or otherwise of translations could be measured. With the development of computing technology able to store and handle massive amounts of linguistic evidence and authentic texts, it has become possible to base linguistic judgment on something far broader and far more varied than any one individual's personal experience or intuitions. "In fact, electronic corpora […] are becoming increasingly used in research across the board in Translation Studies" (Hatim and Munday 2004: 118). This explosive growth is "primarily connected with the possibilities offered by corpora in machine-readable form" (Johansson 2003: 31). Also, the development of the Internet has made the use of electronic corpora not only useful but also indispensable, particularly for evaluators, given that a simple search on the web can yield a wealth of evidence in a matter of just a few seconds that can be used to refute or validate the subjective judgments and intuitive decisions made by the evaluators. It seems that web 2.0 has already left its mark. Those were the days when teachers were the sacrosanct arbiters of language. Now social networking sites like Twitter and Facebook, video-sharing sites, blogs, wikis and the like compensate for the paucity of the very-much-needed native speaker intuition, easily accessible authentic texts, and up-to-the-minute linguistic evidence. However, Internet content

is not necessarily written by language experts or with scholarly rigor, and is certainly not 100% reliable. This is where the immense significance of systematically and methodically developed electronic corpora comes in.

"When training translators, quality assessment should not be an end but a means" (Höning 1998: 32). The same holds true for professional settings since the translation quality of professional translators working for different organizations is sometimes evaluated for different reasons. The reasons can be high pay for high quality, poor pay for poor quality, or as simple as downsizing as a direct consequence of the recent global economic slump that has given new meaning to the law of the survival of the fittest. Whatever the reason, to have better job prospects or to emerge triumphant in the fiercely competitive struggle for survival, professional translators should improve the quality of their translations based on feedback received from evaluations. On the other hand, to establish pay scales or to keep the fittest, employers need to be as fair and objective as possible in their evaluations. This means they also need to have an orientation for translation evaluation. In terms of providing an orientation for translation evaluation, Lauscher (2000: 149) claims that "academic efforts in this area are still largely ignored, if not explicitly rejected by the profession." Moreover, "developing a catalogue of criteria for a 'good' translation" has never been "sufficient for determining translation quality in a professional setting" (Lauscher 2000: 150).

The tangible rewards that can be reaped from basing translation evaluation on electronic corpora in a classroom setting have been shown in Bowker's (2000) seminal article. In fact, the aim of the present paper is to take a step further and see the effects of a corpus-based evaluation approach on translation improvement in a professional setting. This paper also regards translation evaluation only as a means to a worthier end, i.e., any improvement in the quality of translations done by translators. I hypothesize that basing translation evaluation on electronic corpora in a professional setting not only offers a far more objective evaluation by considerably minimizing subjectivity but also helps translators hone their translation skills to relative perfection thanks to objective feedback and convincing corrections constantly received from their evaluators.

The paper consists of four sections. Section 1 focuses on the use of corpora in Translation Studies. Section 2 outlines the advantages of corpora over other conventional resources. Section 3 elaborates on the main features and applications of the Corpus of Contemporary American English (COCA), which has provided insights for the evaluation purposes of this study. Lastly, section 4 reports on a case study carried out to determine the effectiveness and potential consequences of a corpus-based evaluation approach to translation improvement in a professional setting.

A word of notice: translation evaluation, translation criticism, and translation quality assessment have all been interchangeably used with the same sense in the context of this paper.

1. Corpora and Translation Studies

Regarding the huge significance of corpora in Translation Studies, the following claim can be made:

> The information age has brought an explosion in the quantity and quality of information we are expected to master. This, along with the development of electronic modes

> for storing, retrieving, and manipulating that information, means that any discipline wishing to sustain itself in the twenty-first century must adapt its content and methods. Corpus translation studies is central to the way that Translation Studies as a discipline will remain vital and move forward […] corpus translation studies has the characteristics typical of contemporary emerging modes of knowing and investigating (Tymoczko 1998: 652).

It is clear that here Tymoczko means electronic corpora (as opposed to manual corpora) by corpora. In fact, corpora are large collections of texts in electronic format. It is the electronic corpora that are desperately needed for translation purposes since they facilitate the whole process of translation thanks to their information retrieval, storing, and user-friendly facilities. Furthermore, an electronic corpus-based translation is far more reliable because the translator has unrestricted and immediate access to authentic texts and patterns of language use as used by native speakers of the respective languages. Over and above all that, electronic corpora "can be used for purposes which may not have been foreseen at the stage of compilation" (Johansson 1998: 3). A case in point would be the relatively high number of serendipitous findings and incidental revelations in working with corpora (Bowker and Pearson 2002: 200-202).

It is an established fact that bilingual dictionaries are no longer sufficient tools for translators and "totally inadequate as a source of real world data for the purpose of translation studies" (Peters and Picchi 1998: 91). Peters and Picchi (1998: 91) discard them as defective on two basic grounds. First, words do not appear in their natural contexts. The other major problem is the "division of each entry into a set of discrete senses, listing possible translations for each sense." However, corpora can address these shortcomings by providing instant "access to large systematic collections of reliable data attesting the usage of semantically equivalent lexical items across languages, according to variations in style, genre, and text type" (Peters and Picchi 1998: 92).

In recent years there have been frequent recommendations by researchers and trainers alike (Bowker 2000; Laviosa 2003; Frankenberg-Garcia 2009; Bowker and Pearson 2002; Vintar 2008, to name a few) in the field of Translation Studies to integrate the analysis of corpora into translator training. However, although the study of corpora has "become fully integrated into Translation Studies since the early 90s" (Laviosa 2002: 2), only recently has the extensive use of electronic corpora by practising translators gained appropriate ground. The reason for such tardiness is partly due to their lack of exposure to the potential of corpus analysis tools like concordancers and word listers, during their own education, and part in the unavailability of rather comprehensive ready-made special-field corpora.

1.1. Types of Corpora

To use Reppen and Simpson's words (2002: 95), one could claim that "there are as many types of corpora as there are research topics" in Translation Studies. Corpora primarily fall into three general types: monolingual, bilingual, and multilingual corpora. As the names imply, monolingual corpora are comparatively large collections of texts in either source or target language. The texts included in the monolingual corpora are either "texts *originally* produced in a given language" or "texts

translated into that same language from one or more source languages" (Laviosa 1998: 101). The former serve the function of providing naturally occurring patterns of language use and "can provide tremendous insights as to how language use varies in different situations, such as spoken versus written, or formal interactions versus casual conversation" (Reppen and Simpson 2002: 92). However, the latter, which are also technically referred to as monolingual comparable corpora, are used to identify the linguistic nature of translated texts as opposed to that of non-translated texts. As a case in point, one can refer to COCA or to the British National Corpus (BNC) as best examples of monolingual corpora in English containing upwards of 400 and 100 million words respectively. On the other hand, bilingual corpora consist of texts both in the source and target language. Bilingual corpora themselves fall into two types: parallel and comparable corpora. They are discussed in more detail later in this paper. As to multilingual corpora, they comprise source texts with their translations in at least three different languages, e.g., a UN text translated into different languages. Johansson (2003: 40) believes that multilingual corpora, if they truly represent a range of languages, can "increase our knowledge of language-specific, typological, and universal features."

1.1.1. Parallel Corpora

"There is some variation in definition of the term 'parallel corpus' in the literature" (Olohan 2004: 24). Even the naming of 'parallel corpora' has always been a bone of contention so much so that "contrastive linguists refer to [them] as translation corpora" (Granger 2003: 20), causing unnecessary confusion. Annoyingly enough, even contrastive linguists are not totally consistent in their use of the term 'parallel corpus.' Granger (2003: 19) goes so far as to say that contrastive linguists sometimes refer to a parallel corpus as "a comparable corpus (Aijmer *et al.* 1996: 79; Schmied and Schäffler 1996: 41), a translation corpus (Hartmann 1980: 37) or a combined comparable/translation corpus (Johansson *et al.* 1996)." However, the prevailing definition is that parallel corpora are sets of texts in a source language with their corresponding translations in a target language (Peters and Picchi 1998: 92; Ulrych 2002: 207; Olohan 2004: 24). Thus by definition, parallel corpora provide data on translation equivalents. "Parallel corpora can be bilingual or multilingual. They can be unidirectional […], bidirectional […] or multidirectional" (McEnery and Xiao 2008: 20). For instance, if a corpus contains texts from English into Spanish or from Spanish into English alone, it is a unidirectional corpus. However, when a corpus consists of both English source texts with their Spanish translations and the other way round, it is a bidirectional corpus. Finally, if a corpus comprises a source text with its translation in Spanish, Persian, and Portuguese, it is a multidirectional corpus.

Although parallel corpora provide reliable resources for finding translation equivalents, they can have their drawbacks. Granger (2003: 18) believes that as "they often display traces of the source text," they "cannot really be considered as reliable data as regards the target language, especially in frequency terms." Another criticism Granger levels at parallel corpora is the occasional impossibility of finding "translations of all texts, either because of the text type […] or because there are more translations in one direction (English to Norwegian, for instance) than in another (Norwegian to English)."

1.1.2. Comparable Corpora

Comparable corpora "consist of homogeneous sets of texts from pairs of languages which can be contrasted and compared because of their common features" (Peters and Picchi 1998: 92). In fact, comparable corpora "represent natural language use in each language and should allow safe conclusions to be drawn on similarities and differences between the languages compared (Johansson 1998: 5). The texts in comparable corpora have been composed independently in their respective language communities but "are supposed to share certain basic features, such as period of time, topic, functionality, register, domain, etc." (Peters and Picchi 1998: 92). Thus by definition, comparable corpora provide data on natural language lexical items within a given domain (Peters and Picchi 1998: 91).

Because of the self-standing status of texts included in comparable corpora, "they are therefore in principle free from the influence of other languages" (Granger 2003: 19) and can truly represent the languages compared. Granger goes on to say that the major shortcoming of comparable corpora, however, "lies in the difficulty of establishing comparability of texts. Some types of text are culture-specific and simply have no exact equivalent in other languages."

Among the different corpora developed for the express purpose of translation evaluation, Lynne Bowker's (2000) stands out as a pioneering work in the field. She believes that conventional resources available for evaluators, like dictionaries, printed parallel texts, subject field experts and intuition, are inadequate for the purpose of translation evaluation and have their own drawbacks. She also believes that developing an appropriate electronic corpus can help translation evaluators correct these shortcomings.

As regards corpus-based approaches to translation evaluation, Bowker (2000) conducted similar research in an academic setting with some pedagogical implications in mind under the title of "A Corpus-Based Approach to Evaluating Student Translations." Conducting a relatively small-scale experiment to ascertain the helpfulness of a specially designed "Evaluation Corpus" as a resource for evaluators in evaluating student translations, she realized that the test group evaluators, using the "Evaluation Corpus" as their benchmark turned up many more errors than their control group counterparts using conventional resources for the same purpose. Moreover, the test group evaluators' feedback proved considerably more objective than that of the test group evaluators'. Also, the revised texts resulting from the feedback and comments provided by the test group evaluators were judged by both subject field experts and language experts to be of higher quality. Space here does not allow adequate acknowledgement of the detailed presentation of the findings and results of Bowker's study. Anyone with special interest in her study may consult the original paper.

2. Why Electronic Corpora

An electronic corpus is generally a large collection of machine-readable texts that have been gathered according to specific criteria. As Bowker (2000) claims, developing a corpus in electronic form can compensate for the common shortcomings of the conventional resources available to evaluators like printed dictionaries, paper-based parallel texts, subject field experts and intuition. According to Bowker (2000: 186-

190), electronic corpora have some advantages over these conventional resources which are summarized as follows:

2.1. Dictionaries vs electronic corpora

Among the most common problems encountered in using printed dictionaries are lack of space, omission of terms and lack of extended contexts, the relative difficulty of information retrieval, and the substantial amount of time required for publishing and updating printed dictionaries. Also, in this regard, "an additional point of interest […] is that they cannot easily provide information about how frequently a given term is used" (Bowker and Pearson 2002: 16). Bowker (2000: 187) believes that some of these shortcomings are "addressed in dictionaries produced in electronic format." However, even in electronic dictionaries, the bulk of contexts "are often made up by lexicographers" (Bowker 2000: 187) which cannot truly represent the real language in use. Since space is no longer an issue in an electronic corpus, millions of words can easily fit in it. At the same time, running texts in an electronic corpus provide non-lexical information, which does not typically appear in dictionaries. Furthermore, information retrieval is much easier in an electronic corpus thanks to search tools which can facilitate performing some statistical analyses, finding out about the frequency of occurrence of terms in their immediate contexts, and the total number of words used in a corpus, displayed either in an alphabetical or a frequency order. As for recency, some large electronic corpora are updated every year. For instance, COCA is updated every six to nine months, adding about 10 million more words to the corpus in each update.

An additional serious disadvantage of both printed and electronic dictionaries is that "words often are seen as synonymous [like *become*, *turn*, *go* and *come*] when actually, their use is not synonymous" (Reppen and Simpson 2002: 108). However, a corpus search can show the dramatic differences between these words in their appropriate contexts. For instance, for "a change of color," for "a change to a negative state," and for "a change to a more active state," *turn*, *go*, and *come* are used respectively (Reppen and Simpson 2002: 108). Such information is scarce in most dictionaries.

2.2. Parallel texts vs electronic corpora

"Parallel texts are documents that have been produced independently in different languages, but which have the same communicative function as the source text" (Bowker 2000: 188). In the case of parallel texts, if they are paper-based, it is difficult to glean and consult enough documents to make sure that all the relevant concepts and terms exist. Also, simply as a result of human error, spotting linguistic and conceptual patterns is difficult. Gathering these texts in the form of an electronic corpus can help overcome these shortcomings. Corpus analysis tools allow users to conduct more wide-ranging and accurate searches.

2.3. Subject field experts vs electronic corpora

When dictionaries and printed texts fail to provide the information translators or evaluators are trying to obtain, sometimes they may turn to subject field experts for

guidance. However, this approach is not without its setbacks. "First, as observed by Wright (1987: 118), translators are sometimes reluctant to consult subject field experts because they do not wish to appear ignorant" (Bowker 2000: 189). Electronic corpora obviate the problem since one can use them without the embarrassment of appearing ignorant. Second, tight budgets mean that sometimes it does not make economic sense to hire subject field experts. However, one can use corpora free of charge as much and as long as they wish. Third, since experts are human beings, they are not without their own limitations. They may forget something, put it wrongly, or worst of all, express their own views in a prejudiced manner. However, since corpora contain a collection of texts written by different subject field experts, they represent a far broader cross section of the expert views and, therefore, the difficulty can be cleared up. In spite of all this, having pored over related books for years, subject field experts may provide translators and evaluators with a treasure trove of useful resources, a relative rarity in corpora. So I believe that, regardless of all the advantages offered by corpora, there must be a high degree of complementarity between these two resources, and the value of subject field experts should never be underestimated.

2.4. Intuition vs electronic corpora

Being a native speaker of a language does not necessarily mean that one has effortless and absolute mastery of the language, even in translating into one's own mother tongue or evaluating a translation into it, let alone into other languages. Therefore, consulting supplementary resources besides one's own intuition and gut feelings about a language is a *sine qua non*. Relying on one's language intuition to make judgments about which terms or phrases are appropriate can be helpful in the case of LGP (language for general purposes) contexts but not in the case of LSP (language for special purposes) contexts. "Understanding a word's patterns of use is crucial for" both translators and evaluators, and "intuitions often do not prove helpful in predicting patterns" (Reppen and Simpson 2002: 108). However, electronic corpora can be used as a benchmark against which translators or evaluators can verify or refute their intuitions and wild guesses.

3. The Corpus of Contemporary American English

The Corpus of Contemporary American English (COCA), designed in 2008 by Mark Davies,[1] is much used by evaluators in evaluating translators' work as well to provide feedback on their performance, and by translators for translating newspaper texts and improving their translation over a specific period of time in this study.

3.1. Claims and features

According to the designer (Davies), COCA purports to be the largest freely available corpus of English. With upwards of 400 million words of text (at the time of writing this article), it is equally divided among spoken, fiction, popular magazine, newspaper, and academic texts. It includes 20 million words each year from 1990 to 2009 and the corpus is also updated every six to nine months. Its design makes it perhaps the only corpus suitable for looking at current, ongoing changes in the English language.

Exact words or phrases, wildcards, lemmas, parts of speech, or any combination thereof can be searched in this corpus. It is also possible to search for collocates within a ten-word window (e.g., all nouns somewhere near *faint,* all adjectives near *woman*, or all verbs near *feelings*). Limiting searches by frequency and comparing the frequency of words, phrases, and grammatical constructions, synchronically and genre-wise, are a particular boon. Users can also easily carry out semantically-based queries of the corpus. For example, one can compare and contrast the collocates of two related words (*little/small*, *democrats/republicans*, *men/women*), to find the difference in meaning or use between these words. The frequency and distribution of synonyms for nearly 60,000 words can be found and their frequency in different genres can be compared.

3.2. Modus operandi

Davies[2] provides a bird's-eye view of how the corpus works and how it can be worked. Brief as it is, the tour is immensely helpful in guiding users through all the nooks and crannies of the search interface and running a query against the 400-plus million word corpus.

Using the web interface, one can search by words, phrases, lemmas, wildcards, and more complex searches such as *un-x-adjective* or *verb + any word + a form of ground*. It is also possible to see every occurrence in context.

The first option in the search form makes it possible to either see a list of all matching strings, or a chart display showing the frequency in the five *macro* registers (spoken, fiction, popular magazines, newspapers, and academic journals). Via the chart display, one can see the frequency of the word or phrase in subregisters as well, such as movie scripts, children's fiction, women's magazines, or medical journals.

One can also search for collocates. For example, one can search for the most common nouns near *thick*, adjectives near *smile,* nouns after *look into*, or words starting with "*clos** near eyes."

One can also include information about genre or a specific time period directly as part of the query. Users can easily find which words and phrases occur much more frequently in one register than another, such as *good + [noun] in fiction*, or verbs in the slot [*we * that*] *in academic writing*. It is also possible to apply this to collocates, such as *nouns with the verb break* in NEWS or *adjectives with woman in FICTION*. Finally, they can compare one section to another, such as nouns near *chair* (ACADEMIC vs FICTION), nouns with *passionate* (FICTION vs NEWSPAPER), verbs in sports magazines compared to other magazines, or adjectives in medical journals compared to other journals.

Last but not least, carrying out semantically-oriented searches is also possible. For example, one can compare nouns that appear with *few* and *little*, or with *boys* and *girls*, nouns with *utter* and *sheer,* adjectives with *Conservatives* and *Liberals*, or verbs with *Clinton* and *Bush*. Users can also find the frequency and distribution of synonyms of a given word, such as *beautiful* or the verb *clean*, to see which synonyms are more frequent in competing registers (such as synonyms of *strong* in FICTION and ACADEMIC).

3.3. COCA compared to other corpora

Outlining the overall competitive edge of COCA over other corpora freely or conditionally available online like American National Corpus (ANC), British National Corpus (BNC), Bank of English (BoE), and Oxford English Corpus (OEC), Davies enumerates its balance of availability, size, genres, and currency, among others.

The ANC's relatively diminutive size of about 22 million words – in glaring contrast to the whopping 400-plus million words of COCA – and its tardiness in being up-to-date give COCA a commanding lead.

While Davies believes that COCA and BNC complement each other nicely, and they are the only large, well-balanced corpora of English that are publicly available, despite BNC's better coverage of informal, everyday conversation, COCA's quadruple size and surprising recency have important implications for the quantity and quality of the data overall. In the present study, BNC has also been occasionally used to show the difference between the British or American use of some words.

As BoE contains more than 450 million words of both British and American English, it is an amazing corpus but not without its drawbacks. First, unlike COCA, it is not updated any more. Second, the users are charged US $1,150 a year.

Finally, according to Davies, OEC "is a great corpus in terms of its size and even the wide range of genres and text types. COCA is not as large, but it does cover more years." The other disadvantage of OEC is its unavailability to the general public such that very few people can use it.

In Table 1, Davies provides a summary of the features of the different corpora.

Table 1

A summary of the main features of COCA compared to those of BNC, ANC, BoE, and OEC

Feature	COCA	BNC	ANC	BoE	OEC
Availability	free / web	free / web	free	$1150/ year	limited use
Size (millions of words)	400	100	22	455	1,898
Time span	1990-2009	1970s-1993	2000-2005	1970s-2005	2000-2006
Number of words of text being added each year (millions of words)	20	0	0	0	0
Can be used as a monitor corpus	yes	no	no	limited use	limited use
Wide range of genres: spoken, fiction, popular magazine, newspaper, academic	yes	yes	no	yes	yes
Size of spoken (millions of words)	83	10	4	62	82
Spoken = conversational, unscripted?	mostly	yes	yes	some	some
Dialect	American	British	American	American + British	American + British

4. The study

This section sketches out an experiment that was conducted to test and evaluate the usefulness of the translation evaluation approach based on COCA in translation improvement in a professional setting, a news agency in this study. First, twelve

translators working for a news agency in Tehran were selected. They were all university graduates with a B.A. in English Translation or following an M.A. course in Translation Studies. Their translation experience varied from three to seven years. As Persian to English translators are in high demand and paid more generously in Iran, these translators were keen on broadening their scope of translational activities and advancing their professional standing. They were asked to translate a self-contained 283-word news item on international politics, taken from the website of the news agency and belonging to January 2009, from Persian (as their mother tongue) into English (as a foreign language). They were briefed to translate the text as if for publication in an English newspaper such as *Tehran Times*. Then, three evaluators were asked to evaluate these translations according to how well they met the requirement, using conventional resources. When the translations were corrected, errors were identified, comments were provided, and grades were awarded, translators were placed in two groups of six, with better graded ones in the control group and the rest in the test group. Translators in both groups were told to translate six news items like the one administered for placement purposes over a period of twelve weeks, with one every other week, and to try to improve their translation through the comments and feedback they received from their evaluators. From then on, three more evaluators were added to evaluate the translations of the test group. They were told to use COCA as the only available resource for evaluating translations, providing comments and feedback, and identifying and correcting errors as the other three had been told to use only conventional resources for the same task. The evaluators for the test group were also allowed to use BNC only to point out differences between American and British English, if any.

All the evaluators for both the control group and the test group had an M.A. in Translation Studies, had been teaching translation courses at some Iranian universities for five to seven years, and had been working as professional translators for more than ten years. As translator trainers and professional translators, they were all complaining about the difficulty and arbitrariness of translation evaluation both in academic and professional settings. They were all asked to evaluate the translations, firstly, in terms of how well they met the brief and, then, in terms of cohesion, coherence, overall comprehensibility, grammar, register, genre, conceptual understanding, and term choice. The evaluators for the control group did not receive any instructions as to how to evaluate the translations. The test group evaluators, however, received some instructions about the tips and tricks of the corpus, and how to use the corpus to get the intended results.

Translations were evaluated, errors identified and/or corrected, comments and feedback provided, grades awarded and translators' progress was closely observed during that period. Translators were also asked to provide their feedback and express their confidence about the evaluators' comments on translations and their helpfulness in improving their translation skills. The following is an analysis and discussion of the results obtained.

4.1. Translators' predictions about the types and number of errors

In all six translation tests, translators were asked to make a prediction about the types and number of errors they thought they might have made. This was done to see how

aware they were of their own errors and whether this awareness, or lack thereof, would finally yield any improvement in their translation. The evaluated translations more or less confirmed the control group's predictions about both the types and number of errors in all of the tests. With regard to the types of errors, they had mostly predicted mechanical, grammatical, and lexical errors. More interestingly, the evaluators identified no more error types than those predicted by the translators. However, there was a minor discrepancy between their own predictions about the number of errors and the number of errors identified by the evaluators. The evaluators identified roughly one and a half times as many errors as predicted by the translators. This was virtually consistent throughout the all tests. It must also be noted that, as will be seen later, this discrepancy pales in comparison with that of the test group.

On the other hand, with regard to the types of errors predicted by the test group, the range expanded from the initial mechanical, grammatical and lexical errors to more complex errors of collocation, inappropriate linguistic variation (register, genre, dialect, and style), and semantic prosody. As an example of a collocational clash, that is words put together which do not typically go together, two of the test group translators and one of the control group translators had rendered the phrase *in a haste*, for the Persian عجولانه in their second translations. While all three evaluators for the control group had let the error slip by, the other evaluators for the test group had spotted the error and corrected it with compelling evidence from the corpus. According to the comments provided by one of the evaluators, the phrase *in a haste* was used just twice in total in the corpus and both instances can be seen only in spoken English rather than newspaper English. On the other hand, the phrase *in haste* was used 135 times in total, 18 occurrences of which happen to belong to newspaper English. However, his strong recommendation is its more widely used synonym *in a hurry*. The phrase *in a hurry* was used 285 times in the news alone so its much greater frequency seems to be an irresistible temptation. Another exceptionally striking example of a collocational clash can be seen in the translation of the Persian phrase جنايت هاي بي رحمانه in their second test. Four of the control group translators and three of the test group translators suggested *cruel crimes* as an equivalent. One of the control group translators and two of the test group translators suggested *callous crimes*. Also, one of the translators both in the control group and in the test group rendered it as *relentless crimes* and *ruthless crimes* respectively. Evaluated translations showed that while none of the evaluators for the control group had marked any of the renderings as an error, all of the evaluators for the test group had spotted them as errors of collocation by providing ample evidence from the corpus. They all unanimously recommended either *brutal crimes* or *vicious crimes*, prioritizing the former. Commenting on the phrase *cruel crime*, one of the evaluators for the test group dismissed it as a collocational clash and an inappropriate register. Entering the search pattern *[=cruel] [crime]* in the corpus confirms their recommendations as completely valid. Table 2 shows the results produced by entering the pattern.

TABLE 2

Quasi-synonyms of the adjective 'cruel' collocating with the noun 'crime' across different genres

	Collocations	ACAD	NEWS	MAG	FIC	SPOK	TOT
1	*brutal crime*	3	9	4	2	30	48
2	*brutal crimes*	0	3	6	1	7	17
3	*vicious crime*	1	4	2	2	8	17
4	*vicious crimes*	0	4	0	1	3	8
5	*hard crime*	0	1	0	0	1	2
6	*punishing crimes*	2	0	0	0	0	2
7	*unpleasant crime*	0	1	0	0	0	1
8	*ruthless crimes*	0	0	1	0	0	1
9	*painful crime*	0	0	0	0	1	1
10	*nasty crime*	0	0	0	1	0	1
11	*mean crime*	0	0	0	0	1	1
12	*harsh crime*	0	0	0	0	1	1
13	*cruel crime*	0	0	0	0	1	1
14	*callous crimes*	0	0	0	0	1	1
	Total	**6**	**23**	**13**	**7**	**53**	**102**

The abbreviations TOT, SPOK, FIC, MAG, NEWS, and ACAD stand for Total, Spoken, Fictional, Magazine, Newspaper, and Academic, respectively

As can be seen from Table 2, more than half of the collocates listed are one-offs in the corpus. Interestingly enough, five of these one-offs belong to the spoken English, as opposed to the only one occurrence in the newspaper genre. It seems that spoken English is more casual and spontaneous in its approach to the collocates of a specific term in its collocational range.

In respect to the errors of inappropriate linguistic variation, the rendition of تاریکی محض as *stygian darkness* is quite interesting. In their first translation test, one of the control group translators as well as one of the test group translators used the phrase *stygian darkness* as an equivalent for the Persian phrase. While one of the evaluators for the control group had underlined it and praised the translator for his good lexical choice, the other three evaluators for the test group detected it as an error. Feedback from these evaluators shows that the phrase *stygian darkness* was used eight times in total in the corpus: as shown in Table 3, six times in fiction, once in magazines, and only once in newspapers. Its only use in the newspapers dates back to 1996. Since languages are in a constant state of flux, the fact that the phrase has essentially disappeared in this specific genre for the past decade and a half renders it functionally more or less obsolete. Their recommendation is the phrase *total darkness* instead. Evidence from the corpus reveals that the phrase *total darkness* has been used 17 times in newspapers out of a total 183 times. Moreover, it is still in continued use. Table 4 shows the frequency of use of *total darkness* in total along with its frequency in the news in the corpus during four different periods of times.

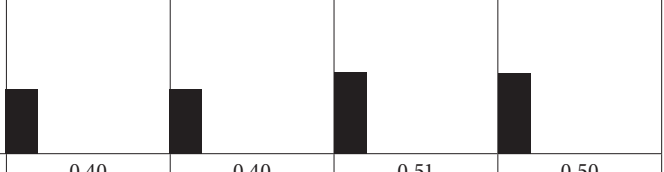
Table 3

Contexts retrieval

	Contexts containing 'Stygian darkness'	Genre of the text	Year of publication
1	…gripped his shoulder. «Now, we wait. «Macromolecules swam purposefully through **stygian darkness**, grazing on glucose dissolved in the murky fluid. The breaking of chemical…	FIC	2009
2	… breathe sweet air again. But in which direction was it? I was in **stygian darkness**, and my convulsions had even lost me the bottom I had just rested…	FIC	2006
3	… THE END 828 «THE COOLER «by Frank Hannah and Wayne Kramer EXT. **STYGIAN DARKNESS** - NIGHT STYGIAN DARKNESS The suggestion of traveling through space. Suddenly a star…	FIC	2003
4	… THE COOLER «by Frank Hannah and Wayne Kramer EXT. STYGIAN DARKNESS - NIGHT **STYGIAN DARKNESS** The suggestion of traveling through space. Suddenly a star sparkles to life in…	FIC	2003
5	… to pieces, and then the flashlight goes out and again, we're in **stygian darkness**. The sound of Bill dying and a long silence. PATRICK (CONT…	FIC	2001
6	… had survived as well. He emerged from the World Financial Center garage into the **stygian darkness** of dust and smoke. The first deputy commissioner immediately began digging out bodies	MAG	2001
7	…, and for a moment they were plunged into what seemed to be absolute, **stygian darkness**. Then, after a few agonizing seconds, a faint, yellowish-green outline…	FIC	1999
8	… builds in layers and intensity, its luminescent aural hues painted across a backdrop of **stygian darkness**. # Part's music is not for type A personalities as it unfolds…	NEWS	1996

Table 4

Frequency of use: diachronic and genre analysis

Years	2005-2009	2000-2004	1995-1999	1990-1994
				
Per Million	0.40	0.40	0.51	0.50
Total Frequency	37	41	53	52
Frequency in Newspapers	4	1	5	7

The frequency of use of the phrase 'total darkness' as given for the whole corpus and in newspapers for four different periods of time.

Errors of semantic prosody were also detected in the translations provided by the test group. Semantic prosody is "the spreading of connotational colouring beyond single word boundaries" (Partington 1998: 68). Connotation refers to the "attitudes and emotions" reflected by words that can be positive, negative, or neutral (Larson 1984: 131). In the fourth translation test, the Persian text contained the phrase بیانیه ای مالامال از تشویش. The adjective *مالامال*, a formal word generally meaning *full of*, collocates with both positive and negative words, in this case with a negative word

meaning *anxiety*. Probably consulting a bilingual Persian-to-English dictionary, two of the translators in the test group and four of the translators in the control group translated it into *brimful of*. While only one of the evaluators for the control group had underlined it as unnatural English, two of the evaluators for the test group identified the translation as an error of semantic prosody. Evidence from the corpus shows that the English phrase *brimful of* collocates with words like *confidence*, *vigour*, *flowers*, *sun*, and *visions*. It never collocates with a negative word like *anxiety*. Instead the phrase *fraught with* has a negative semantic prosody which can collocate with words like *danger*, *difficulties*, *anxiety*, *tension*, *uncertainty*, *peril*, and *problems*.

As for the number of errors predicted by the translators, since the translators in the test group received more convincing and corrective feedback on a wider range of error types, they became progressively more aware of and sensitized to the types of errors they might have made, and this in turn led to a progressive reduction in the number of errors made from the first to the last test. Common sense dictates that when you are not aware of an error, you make no attempt to correct it. While the number of errors predicted by them increased rather significantly from test to test, the number of errors identified by the evaluators progressively closed in on the translators' predictions. For instance, as these translators were not familiar with a wider range of errors in the first tests, the errors identified in their first test were roughly four times as many as those of their predictions. Table 5 shows the number of errors predicted by the translators in both groups and the average number of errors identified by the three evaluators for each group.

A closer look at the numbers in Table 5 reveals that while there is only a slight difference between the number of errors predicted by the control group translators and identified by their evaluators, there is a yawning gap in the number of errors predicted by the test group translators and identified by the relative evaluators and this continues to narrow. This again suggests that the evaluators for the control group did nothing special to raise the awareness of the translators to a wider range of errors and finally to help them lower the number of errors they made from test to test as the number of identified errors remained almost consistent throughout the whole set of tests. Nevertheless, with a growing awareness to a wider range of errors, test group translators were able to predict more errors and lower the number of errors made from test to test. It is tempting to speculate that the control group translators were much better translators than their test group counterparts. That is why the number of errors predicted by them and identified by their evaluators were far less. To begin with, they were better translators. This claim was substantiated by the placement test and the language experts' views on their first test which is discussed in detail in the following sections. Having started as better translators never means that they ended up as better ones as well. Additionally, the cross-evaluation of the final translations showed that the control group translators did not end up the same. This is also discussed in detail in the following sections.

Table 5

Number of errors predicted by each translator and average number of errors identified by the three evaluators in each group

	1st Test	2nd Test	3rd Test	4th Test	5th Test	6th Test
Control Group						
Translator A	8 (11)	9 (13)	6 (9)	8 (10)	10 (14)	7 (11)
Translator B	10 (14)	12 (17)	11 (15)	10 (13)	8 (10)	9 (12)
Translator C	12 (17)	10 (12)	10 (13)	14 (19)	10 (13)	8 (11)
Translator D	3 (4)	5 (7)	4 (6)	3 (5)	6 (8)	4 (6)
Translator E	13 (18)	15 (21)	12 (17)	11 (15)	8 (9)	8 (14)
Translator F	10 (15)	10 (13)	8 (11)	7 (10)	12 (14)	10 (14)
Test Group						
Translator G	15 (57)	19 (50)	28 (43)	30 (40)	31 (42)	35 (39)
Translator H	17 (65)	20 (58)	22 (60)	25 (55)	30 (47)	33 (40)
Translator I	18 (70)	22 (63)	28 (57)	31 (50)	35 (48)	40 (44)
Translator J	17 (59)	21 (50)	25 (48)	33 (50)	35 (46)	37 (43)
Translator K	16 (55)	19 (51)	23 (48)	28 (43)	32 (40)	36 (38)
Translator L	17 (49)	19 (44)	22 (41)	28 (42)	32 (38)	34 (35)

The first number in each cell shows the number of errors predicted by each translator, while the number in parentheses shows the average number of errors identified by the three evaluators in each group.

Table 6 shows the discrepancy between the number of errors predicted by the translators in both groups and identified by their evaluators in the first and the last tests. The data indicate that the greatest discrepancies in the control group in the number of errors predicted and identified are minus five and minus six respectively. Table 6 also shows that in the most special case, in the last test, one of the translators predicted only one error more than the number of errors predicted in the first test and also another translator's number of identified errors was six errors less than the number identified in the first test. On the other hand, in the most striking case, one translator in the test group predicted 22 more errors in the last test than his predicted number in the first. Moreover, in this regard, the same translator's number of identified errors was 26 errors less than the number identified by the evaluators in the first test. This finding suggests great progress and improvement from the first to the last translation. In total, test group translators' number of predicted errors increased significantly, whereas the number of their identified errors decreased substantially.

Table 6

Discrepancy between the number of errors predicted by the translators in both groups and the average number of errors identified by the evaluators in their first and last translations

Control Group		Test Group	
Translator A	-1 (0)	Translator G	+20 (-18)
Translator B	-1 (-2)	Translator H	+16 (-25)
Translator C	-4 (-6)	Translator I	+22 (-26)
Translator D	+1 (+2)	Translator J	+20 (-16)
Translator E	-5 (-4)	Translator K	+20 (-7)
Translator F	0 (-1)	Translator L	+17 (-14)
Total	**-10 (-11)**	**Total**	**+115 (-116)**

The first number in each cell shows the predicted number while the number in parentheses shows the average number of errors identified by the evaluators.

In this regard, Table 7 also shows the total number of errors predicted by the translators and identified by the evaluators in all six translations for both groups. As the data in Table 7 suggest, the evaluators for the test group identified about four times as many errors as their counterparts for the control group did.

To sum up this rather long discussion, the wider range of errors and the more errors an evaluator identifies, the wider range of errors and the more errors a translator is likely to predict in subsequent translations. This in turn leads to a progressive reduction in the number of errors made by the translator as she/he becomes more aware of the different types of errors and tries to avoid them in subsequent translations. The opposite also seems to hold true.

TABLE 7

Total number of errors predicted by the translators and identified by the evaluators in all tests for both groups

Control Group		**Test Group**	
Translator A	48 (68)	Translator G	158 (271)
Translator B	60 (81)	Translator H	147 (325)
Translator C	64 (85)	Translator I	174 (332)
Translator D	25 (36)	Translator J	168 (296)
Translator E	67 (94)	Translator K	154 (275)
Translator F	57 (77)	Translator L	152 (249)
Total	**321 (441)**	**Total**	**953 (1748)**

The first number in each cell shows the total number of errors predicted while the number in parentheses shows the average total number of errors identified by the evaluators.

4.2. Testing translators' progress

To see if feedback provided and corrections made through evaluations had any effect on the overall quality of translations and made the translators any more competent and skilled in the end, their final translations were cross-evaluated and expert advice was sought. As an additional test of progress, the translators' avoidance of a repeated error was observed.

4.2.1. Cross-evaluation of the final translations

After being originally evaluated by the relative evaluators, the final translations of both groups were passed on to the evaluators of the opposite group. This was done to see if the far greater number of errors identified by the evaluators for the test group in comparison with the far fewer number of errors identified by the evaluators for the control group was a matter of personal taste or a product of the approach they had adopted in evaluating the translations. The results are presented in Table 8.

TABLE 8

The average number of errors identified for each individual translator in the final test in cross-evaluations.

Control Group		Test Group	
Translator A	52 (11)	Translator G	10 (39)
Translator B	48 (12)	Translator H	7 (40)
Translator C	42 (11)	Translator I	8 (44)
Translator D	47 (6)	Translator J	8 (43)
Translator E	60 (14)	Translator K	4 (38)
Translator F	58 (14)	Translator L	8 (35)
Total	**307 (68)**	**Total**	**45 (239)**

The first number in each cell shows the average number of errors identified in cross-evaluation while the number in parentheses shows the average number of errors identified by the relative evaluators for each group.

Interestingly enough, cross-evaluations revealed a fewer number of errors for the test group translators and a much greater number for the control group translators. This finding highlights two points. First, the greater number of errors identified for the test group or the fewer number of errors identified for the control group throughout the tests was less a product of a harsh or laissez-faire attitude taken by the evaluators than a result of the evaluation approach they had adopted. In fact, the conventional resources the evaluators for the control group had at their disposal did not give them the essential equipment to spot a broader range and greater number of errors. On the other hand, the bleeding-edge technology of the corpus equipped the evaluators for the test group with a value-priced wealth of information on language use, authentic patterns of language, information retrieval capabilities, and abundant representative examples from as varied a genre as fiction, spoken, academic, magazine, and newspaper so that they could match the offered equivalents with the real ones as used by native speakers and mark the errors on the spot.

Second, it suggests the overall superior quality of the test group's final translations and a marked improvement in their translations during this period of time from the first to the last translation. Although the relatively weaker translators were put in the test group, they ended up better translators than their control group counterparts. This steady and encouraging progress can be ascribed to the corpus-based evaluation of their translations, objective feedback and constructive comments provided on their translations, and corrections made on their errors based on tangible evidence rather than the subjective judgments of the evaluators. Since none of the translators in either group had been asked or advised to use the corpus for their translation, the effect of a corpus-based translation on the overall quality of translations certainly warrants further study.

4.2.2. Experts' views on the quality of the translations

To see any improvement in the quality of the translations, I also asked six news editors in the news agency to rank the first and the last translations of both groups from best to worst, with 1 meaning the best and 12 meaning the worst. These editors were primarily language experts who had become subject field experts with years of experience in the news agency. In fact, they determine the quality of translations done by translators and report it to the employer as part of their job description. They were

asked to rank the quality of the translations based on their suitability for publication in an English language newspaper such as *Tehran Times*, i.e., the translators' brief.

As shown in Table 9, all the top three positions went to the translations done by the control group in the first test, while all the bottom three positions were given to the translations done by the test group in their first translations. Moreover, all the experts gave the first top positions in their ranking order to a translation done by one of the control group members. This ranking proves the overall higher quality of the translations done by the control group in the first test before receiving any feedback on their quality by the evaluators.

TABLE 9

Rankings of the first translations as given by experts

	Expert 1	Expert 2	Expert 3	Expert 4	Expert 5	Expert 6
Control Group						
Translation A	2	3	3	2	3	2
Translation B	5	4	5	3	2	3
Translation C	6	9	4	8	7	6
Translation D	1	1	1	1	1	1
Translation E	6	4	6	7	5	7
Translation F	4	3	2	2	4	5
Test Group						
Translation G	6	8	5	9	7	11
Translation H	12	10	11	8	10	12
Translation I	7	12	10	9	12	11
Translation J	12	10	8	11	10	9
Translation K	9	11	12	10	11	9
Translation L	4	8	7	5	8	6

The results as to the overall quality of the last translations are a far cry from those of the first translations. As the data in Table 10 illustrate, except on one occasion, experts gave the first, second, and third top positions to the translations done by the test group. The translation in the control group that got the first three top positions only once turns out to be the one that got all the first top positions in the first translation. Furthermore, none of the translations done by the test group received the ranking of 11 or 12, and only one of the translations of the test group got the ranking of 10. This finding is another proof of improvement in the overall quality of the translations that had received feedback and corrections based on the corpus.

TABLE 10

Rankings of the last translations as given by experts

	Expert 1	Expert 2	Expert 3	Expert 4	Expert 5	Expert 6
Control Group						
Translation A	4	6	12	12	5	8
Translation B	10	6	6	10	5	7
Translation C	10	11	11	12	9	12
Translation D	2	5	3	5	7	1
Translation E	9	11	10	7	11	12
Translation F	11	11	6	10	7	12

Test Group						
Translation G	2	3	1	3	8	1
Translation H	2	7	3	9	3	2
Translation I	9	4	8	6	9	5
Translation J	10	4	5	7	8	9
Translation K	1	8	3	4	8	2
Translation L	6	4	2	4	1	1

4.2.3 Avoidance of a repeated error

To see how the corrections made on translation errors raised the translators' awareness to avoid making the same errors in subsequent translations, a phrase that had been mistranslated in all the translations of both groups in the second test was picked and used in the fifth one. انتخابات رياست جمهوري ۲۰۰۴ is the phrase that was picked.

Since this phrase is always used in the plural in Persian, all the translators had used the plural form *2004 presidential elections* in English as well. It is clear that the translators had made the error for failing to accommodate the intra-system shift as discussed by Catford (1965/2000: 146), that is, the existence of approximately corresponding systems within two languages but selecting *a non-corresponding term* as a translation requirement. For example, although English and Persian have similar systems of number in terms of being singular or plural – e.g., Eng. *the book/the books* = Per. کتاب/کتاب‌ها –, "in translation, however, it quite frequently happens that this formal correspondence is departed" (Hatim and Munday 2004: 146). As a case in point, the Persian اخبار which is plural, becomes *news* in English. Four of the evaluators for the test group and only one of the evaluators for the control group had spotted the error. While the evaluator for the control group had only crossed out the plural '*s*' and dismissed it as unnatural English, the other four evaluators for the test group had beefed up their arguments with incontrovertible evidence from the corpus. According to the explanation provided by one of the evaluators, the English phrase *presidential election* and *presidential elections* have been used 640 and 221 times respectively in the newspaper genre in the corpus in total. The evaluator also points out that when these phrases are confined to a specific year like *2004*, *2000*, or *1996 presidential election(s)*, the plural frequency suffers disproportionately. For instance, the corpus shows that while the phrase *2000 presidential election* was used 45 times in the newspaper genre, the *2000 presidential elections* was used only twice. Convincingly enough, while the phrase *1996 presidential elections* was used only twice, as opposed to the ten time occurrence of the *1996 presidential election*, on one occasion it is the context, not the built-in feature of the phrase, that dictates the plural '*s*' to it. The context in which it was used is *in the 1992 and 1996 presidential elections*.

Table 11 illustrates that when that phrase was used again in another test, none of the test group translators repeated the same translation error and only one of the control group translators avoided the error and the other five made the same error. This finding also backs up the claim that since corpora provide authentic examples and convincing evidence, corrections based on corpora can help translators a great deal to have a better understanding of the nature of their errors, to realize where and why they went wrong, and, as a result, to improve their translation skills.

TABLE 11

Avoidance of an error that was made in a previous test

Control Group		Test Group	
Translator A	no	Translator G	yes
Translator B	no	Translator H	yes
Translator C	no	Translator I	yes
Translator D	yes	Translator J	yes
Translator E	no	Translator K	yes
Translator F	no	Translator L	yes

4.3. Other findings

Space here does not allow justice to be done to a detailed discussion of all other findings of the research. However, I shall touch on a few other findings that are noteworthy. First, I asked the translators to briefly comment on each of the evaluators' overall performance and to grade the usefulness of the feedback they received from each of the evaluators on a scale from 0 to 100. On average, the evaluators for the control group and the test group scored 35 and 90 respectively. The control group translators mostly cited the high subjectivity of the feedback and the lack of consistency in the identification of errors by all of the evaluators as severe limitations. One control group translator interestingly mentioned that results produced by the search giant Google and the corrections made by the evaluators sometimes contradicted each other. To support his claim, he pointed out that, much to his bewilderment, one of the evaluators had crossed out the plural '*s*' in the phrase *presidential elections* while a look-up on Google yielded myriad examples of the phrase in the plural. That is where the corpus comes in. Although Google or the Web in general, is much larger than any corpus, "using Google as a full-blown linguistic search engine has real drawbacks" (Davies). Limitations on searching for differences in style, finding out about language changes over time, and doing semantically or grammar-based searches are just a few demerits of Google (Davies). Conversely, the test group translators enumerated the objectivity of the feedback, the authenticity and abundance of the examples provided by their relative evaluators as some of the strengths of their approach. They also pointed out that they were able to improve their translations as they became familiar with a broader range of errors, a greater number of errors were identified in their translations, and documented evidence from the corpus left almost no room for anecdotal evidence, intuition and idiosyncratic preferences.

Additionally, after the fashion of Bowker (2000: 195-198), the relativity of the number of errors identified and corrected by the evaluators of both groups were examined and the number of subjective statements used in feedback by them were counted. The findings bore out her observations that the evaluators for the test group not only identified more errors but corrected a higher number of errors, and they also used much fewer subjective comments than the evaluators for the control group did.

5. Conclusion

In a professional setting, translation evaluation is necessary "for the sake of the client's peace of mind" and "the practitioner's professional standing" (Shuttleworth

1998: 78). Drawing on a case study, in this paper I have argued that a corpus-based translation evaluation can benefit both clients and translators. Thanks to the large volume of data stored in electronic format, authentic examples, useful information on style and language changes over time, the possibility of doing a search on the standing and behaviour of a given word as to other words, and the easy retrieval of data, corpora can help both evaluators to be less subjective in their evaluation and translators to learn from feedback in order to improve their skills and enhance their standing. However, if the notion of corpora intrigues you, just a caveat is warranted here: corpora are no magic bullets but alleviators. They only provide evaluators as well as translators with a parallax view not available in conventional resources. Subjectivity is tightly woven into the warp and weft of the translation evaluation fabric. Every attempt should be made to ward it off to the best of our ability. Finally, despite the ever-growing interest in machine translation, translation is still largely a human activity. So, too, is its evaluation. No machine or software, however colossal, can ever completely supplant the human judgement.

NOTES

1. COCA is freely available at <http://www.americancorpus.org>, visited on 28 March, 2010.
2. Davies, Mark (2008-): *The Corpus of Contemporary American English (COCA): 400+ million words, 1990-present*. Visited on 28 March, 2010, <http://www.americancorpus.org>. All the information related to this corpus can be found via the online help on the home page.

ACKNOWLEGMENTS

I would like to thank all the translators, evaluators and language and subject field experts who spared no efforts to help me conduct the research. A word of thanks as well to Lynne Bowker, whose work on translation evaluation continued to be inspirational throughout.

REFERENCES

Bowker, Lynne (2000): A Corpus-Based Approach to Evaluating Student Translations. *The Translator*. 6(2):183-210.

Bowker, Lynne (2001): Towards a Methodology for a Corpus-Based Approach to Translation Evaluation. *Meta*. 46(2):345-364.

Bowker, Lynne and Pearson, Jennifer (2002): *Working with Specialized Language: A practical guide to using corpora*. London: Routledge.

Catford, John C. (1965/2000): *A Linguistic Theory of Translation: An essay in applied linguistics*. London: Oxford University Press (1965).

Farahzad, Farzaneh (2003): A Theoretical Framework for Translation Criticism. *Translation Studies*. 1(3):29-36.

Frankenberg-Garcia, Anna (2009): Are translations longer than source texts? A corpus-based study of explicitation. *In*: Allison Beeby, Patricia Rodríguez Inés and Pilar Sánchez-Gijón, eds. *Corpus Use and Translating*. Amsterdam and Philadelphia: John Benjamins, 47-58.

Granger, Sylviane (2003): The corpus approach: a common way forward for Contrastive Linguistics and Translation Studies? *In*: Sylviane Granger, Jacques Lerot and Stephanie Petch-Tyson, eds. *Corpus-based Approaches to Contrastive Linguistics and Translation Studies*. Amsterdam and New York: Rodopi, 17-29.

Hatim, Basil and Munday, Jeremy (2004): *Translation: An Advanced Resource Book*. London and New York: Routledge.

Höning, Hans G. (1998): Positions, Power and Practice: Functionalist Approaches and Translation Quality Assessment. *In*: Christina Schäffner, ed. *Translation and Quality*. Clevedon: Multilingual Matters, 6-34.

House, Juliane (1981): *A Model for Translation Quality Assessment*. Tübingen: Gunter Narr Verlag.

House, Juliane (1993): The Evaluation of Translation. *In*: Ron Asher, ed. *The Encyclopedia of Language and Linguistics*. London: Pergamon Press, 4700-4708.

House, Juliane (1998): Quality of Translation. *In*: Mona Baker, ed. *The Routledge Encyclopedia of Translation Studies*. London and New York: Routledge, 197-200.

Johansson, Stig (1998): On the role of corpora in cross-linguistic research. *In*: Stig Johansson and Signe Oksefjell, eds. *Corpora and Cross-linguistic Research: Theory, Method and Case Studies*. Amsterdam and New York: Rodopi, 3-24.

Johansson, Stig (2003): Contrastive linguistics and corpora. *In*: Sylviane Granger, Jacques Lerot and Stephanie Petch-Tyson, eds. *Corpus-based Approaches to Contrastive Linguistics and Translation Studies*. Amsterdam and New York: Rodopi, 31-44.

Larson, Mildred L. (1984): *Meaning-based Translation: A Guide to Cross-language Equivalence*. Lanham: University Press of America.

Lauscher, Susanne (2000): Translation Quality Assessment: Where Can Theory and Practice Meet? *The Translator*. 6(2):149-168.

Laviosa, Sara (1998): The English Comparable Corpus (ECC): A Resource and a Methodology for the Empirical Study of Translation. *In*: Lynne Bowker, Michael Cronin, Dorothy Kenny, *et al.* eds. *Unity in Diversity? Current Trends in Translation Studies*. Manchester: St. Jerome, 149-165.

Laviosa, Sara (2002): *Corpus-based Translation Studies: Theory, Findings, Applications*. Amsterdam and New York: Rodopi.

Laviosa, Sara (2003): Corpora and Translation Studies. *In*: Sylviane Granger, Jacques Lerot and Stephanie Petch-Tyson, eds. *Corpus-based Approaches to Contrastive Linguistics and Translation Studies*. Amsterdam and New York: Rodopi, 45-54.

Maier, Carol (2000): Introduction. *The Translator*. 6(2):137-148.

McEnery, Tony and Xiao, Richard (2008): Parallel and Comparable Corpora: What is Happening? *In*: Gunilla Anderman and Margaret Rogers, eds. *Incorporating Corpora: The Linguist and The Translator*. Clevedon: Multilingual Matters, 18-31.

Olohan, Maeve (2004): *Introducing Corpora in Translation Studies*. New York: Routledge.

Partington, Alan (1998): *Patterns and Meanings: Using Corpora for English Language Research and Teaching*. Amsterdam and Philadelphia: John Benjamins.

Peters, Carol and Picchi, Eugenio (1998): Bilingual Reference Corpora for Translators and Translation Studies. *In*: Lynne Bowker, Michael Cronin, Dorothy Kenny, *et al.* eds. *Unity in Diversity? Current Trends in Translation Studies*. Manchester: St. Jerome, 91-100.

Reiss, Katharina (1971/2000): *Translation Criticism – The Potentials and Limitations: Categories and Criteria for Translation Quality Assessment*. (Translated by Erroll F. Rhodes) Manchester: St. Jerome.

Reppen, Randi and Simpson, Rita (2002): Corpus Linguistics. *In*: Norbert Schmitt, ed. *An Introduction to Applied Linguistics*. New York: Arnold, 92-111.

Shuttleworth, Mark (1998): Preparing Professionals: A Response to Hans G. Höning. *In*: Christina Schäffner, ed. *Translation and Quality*. Clevedon: Multilingual Matters, 78-82.

Tymoczko, Maria (1998): Computerized Corpora and the Future of Translation Studies. *Meta*. 43(4): 652-659.

Ulrych, Margherita (2002): An evidence-based approach to applied translation studies. *In*: Alessandra Riccardi, ed. *Translation Studies: Perspectives on an Emerging Discipline*. Cambridge: Cambridge University Press, 198-213.

Vintar, Špela (2008): Corpora in Translator Training and Practice: A Slovene Perspective. *In*: Gunilla Anderman and Margaret Rogers, eds. *Incorporating Corpora: The Linguist and The Translator*. Clevedon: Multilingual Matters, 153-167.