

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/308210014>

The Application of Corpora in Translation Teaching: A Critical Review

Article · December 2014

CITATIONS

2

READS

1,632

1 author:



[Kanglong Liu](#)

The Hong Kong Polytechnic University

74 PUBLICATIONS 323 CITATIONS

[SEE PROFILE](#)

Translation Quarterly

No. 74 2014

Special Issue:

Intersections and

Interconnections in

Translation Research (III)

香港翻譯學會出版

翻譯季刊

二〇一四年
第七十四期

Published by
The Hong Kong Translation Society

《翻譯季刊》

二〇一四年十二月 第七十四期

版權所有，未經許可，不得轉載。

Translation Quarterly

No. 74, December 2014

All Rights Reserved

Copyright © 2014 THE HONG KONG TRANSLATION SOCIETY

ISSN 1027-8559-74



The Hong Kong Translation Society has entered into an electronic licensing relationship with EBSCO Publishing, the world's most prolific aggregator of full text journals, magazines and other sources. The full text of the *Translation Quarterly* can be found on EBSCO Publishing's databases.



目錄 CONTENTS

論文 Articles

- | | | |
|----|--|--------------|
| 1 | 《哈姆雷特》散文體和詩體漢譯本辭
格翻譯對比研究 | 謝桂霞 |
| 36 | The Application of Corpora in Translation
Teaching: A Critical Review | Liu Kanglong |
| 70 | 從三部作品看方言文學在中國的翻譯
和出版 | 余靜 |

研究札記 Research Notes

- | | | |
|-----|--|------------|
| 93 | Legal Translation: A Lawyer's Perspective | Foster Yim |
| 102 | 稿約凡例 Guidelines for Contributors | |
| 106 | 徵求訂戶啟事 Subscribing to <i>Translation
Quarterly</i> | |
| 108 | 訂戶表格 Subscription and Order Form | |

The Application of Corpora in Translation Teaching: A Critical Review

Liu Kanglong

Abstract

As suggested by its name, corpus-assisted translation pedagogy is placed within a triangle formed by at least three distinct but not discrete disciplines: corpus linguistics, translation and pedagogy. The interdisciplinary nature of corpus-assisted approaches to translation teaching more or less dictates that their investigation should take into consideration the influences and theories of these three areas. In the past two decades, there has been an exponential increase in research studies advocating the adoption of corpora for translation teaching. This paper brings together various strands of research and development in the field of corpus-assisted translation teaching over the past two decades. It begins with an overview of corpus applications in translation research, followed by a critical review of recent research dealing with: (1) corpus-based studies in language pedagogy; (2) different types of corpora used in translation teaching; and (3) issues and debates in the application of corpora to translation teaching. The paper concludes with a discussion of areas for further research.

1. Corpus Linguistics and Translation Research

It is generally agreed among translation researchers that Baker (1993)

is the first scholar to envisage and propose a corpus-based approach to researching translation. In a number of her papers (1993, 1995, 1996), she advances the methodology for corpus-based research and examined the possibility of using such a methodology to investigate authentic linguistic data. In her seminal paper “Corpus Linguistics and Translation Studies: Implications and Applications”, Baker (1993) predicts that the availability of large corpora together with the application of corpus-based methodologies in Translation Studies would “have a direct impact on the emerging discipline of translation studies, particularly with respect to its theoretical and descriptive branches” (233). She recommended that translation scholars use corpus techniques to investigate “the nature of translated texts as a mediated communicative event” (243).

In fact, researchers had been using corpus-like methodologies to investigate linguistic phenomena long before the advent of modern corpus linguistics (cf. McEnery & Wilson 2001). What should be kept in mind is that the rapid development of computer technology and availability of corpus software in recent years have enabled corpus linguistics to develop into a full-fledged methodology. The situation is the same in the field of translation studies. Before Baker put forward a corpus-based approach to translation studies, researchers had identified some unique linguistic features of translational language through contrastive studies of translated texts vis-à-vis their correspondent source texts (e.g. Blum-Kulka & Levenston 1983; Vanderauwera 1985; Shlesinger 1989, 1991). The methodology employed by these researchers is indeed corpus-like, if not strictly corpus-based, considering that their corpus is paper-based and that they had to manually read over the texts to analyze data and calculate frequencies. The basic assumption of this type of research is that translation is fundamentally a mediating process and as a recodification it has its own unique features which might come under the interference of the source text (Toury 1995: 274-279). For example, Frawley (1984: 167-168) argues that “translation is

the bilateral accommodation of a matrix and target code...The translation itself, as a matter of fact, is essentially a third code which arises out of the bilateral consideration of the matrix and target codes; it is, in a sense, a subcode of each of the codes involved.”

Based on these previous assumptions and research findings made by researchers using a corpus-like approach, Baker (1993) reformulates the hypothesis and proposes the concept of “universal features of translation” (which was later termed as “translation universals”) for corpus-based investigations. According to Baker (1993: 243), universal features of translation are “features which typically occur in translated texts rather than original utterances and which are not the result of interference from specific linguistic systems”. The role “interference” plays in forming the features of translational language is a controversial one, and researchers such as Baker (1993) and Toury (1995) hold different views regarding this issue. In reformulating the above hypotheses, Baker (1993) did not include the concept of “interference” while Toury (1995), when proposing a similar concept of “translation laws”, suggested the “law of interference”, according to which “phenomena pertaining to the make-up of the source text tend to be transferred to the target text” (1995: 275). Although their views on this issue differ, the use of corpora for investigating translation phenomena proposed by Baker is also welcomed by Toury, who states that “the quest for [translational] laws would have to take into full consideration regularities of actual behavior obtained by an evergrowing (and ever more variegated) series of studies into well-defined corpuses” (1995: 265). The translation universals proposed by Baker (1993) include simplification, explicitation, normalization or conservatism, and leveling out. Since then, researchers have used corpus-based methods to conduct research to uncover the potential universals of translational language. Although the use of corpora for studying translation universals, as proposed by Baker, is not a novel thing, the approach of studying translated texts through their comparison with non-translated texts^[1] in the same language instead of their correspondent source texts is methodologically rather innovative. Baker’s approach has greatly

elevated the status of translation in the sense that translation research does not necessarily have to be conducted based on the traditional concept of equivalence, which presupposes the underlying relationship between source texts and target texts. Translated texts, which are traditionally treated as a derivative secondary product, are studied in their own right within the target socio-cultural system. This is in line with the development of Descriptive Translation Studies, which also emphasizes a target-oriented approach.

Since Baker put forward these hypotheses, the number of corpus-based studies has increased considerably and most of this research aims at studying translation universals as proposed by Baker. The types of corpora related to translation research include the monolingual comparable corpus^[2] (e.g. Laviosa 1998; Olohan & Baker 2000; Olohan 2003, 2004) and the parallel corpus (e.g. Øverås 1998; Kenny 2001). The translation universals examined using corpus linguistics include simplification (e.g. Laviosa 1998), explicitation (e.g. Øverås 1998; Olohan & Baker 2000), normalization (e.g. Kenny 2001) and the Unique Items Hypothesis proposed by Tirkkonen-Condit (2004), which is also studied and confirmed by Eskola (2004).

It is not the aim of this paper to review every corpus-related translation study. Rather, the studies are mentioned here because they represent pioneering research in corpus-based translation studies (CTS) and they have, to a certain extent, informed and inspired corpus-assisted translation teaching from a methodological perspective. The systematic description of the nature of translated texts is sure to inform translation pedagogy in general, as can be substantiated by Toury's Conceptual Map of Translation Studies (1995).

In Toury's Map of Translation Studies, the applied branch (e.g. translator training, translation aids, translation criticism or translation planning) is under the influence of theoretical assumptions, predictions and even findings concerning translational behavior (Toury 1995: 17-18). In this sense, corpus-based research into translation universals is closely related to teaching. As argued by Laviosa (2008):

Corpus-based research into the universals of translation is strengthening the pivotal role of description in Translation Studies through the development of an explicit, coherent methodology and the acquisition of new knowledge about translational behavior, without necessarily paying attention to such bridging rules. At the same time, however, teachers of translation are independently drawing on the insights of corpus-based Descriptive Translation Studies and would seem to be aiming, in the long term, to formulate bridging rules that postulate what translator trainees should be doing if they adhere to the patterns of translational behaviour unveiled by descriptive scholars. (119)

In many ways, in the field of translation studies, the rise of corpus linguistics is inextricably aligned with DTS, which “foregrounded the description of what translation was and is, removing from dominance previous approaches that were more concerned with prescribing what translation should be” (Olohan 2004: 10). In this way, the translated texts are given a central instead of a peripheral role in the target socio-cultural system. This means that they do not necessarily have to be studied alongside the source texts. Instead, they can be studied using the comparative method of analyzing the translated texts as opposed to the non-translated texts in the same language. It is because of this that DTS is making its way into CTS as the two share a number of similar tenets. Laviosa (2008: 122) even coins the term “corpus-based descriptive translation studies” to emphasize the relationship between CTS and DTS. Laviosa (*ibid.*) states that “electronic corpora have been widely used as a research methodology in Descriptive Translation Studies”. The conceptualization of corpus methodology as a sub-branch of DTS by mainstream CTS scholars clearly indicates the close association between the two approaches.

From the pedagogical perspective, the descriptive nature of corpora has implications for translation teaching as well. Traditionally, translation teaching is perceived largely as a prescriptive activity that allows little room for descriptive interpretation. Hence, translation teaching, which falls under

the category of Applied Extension of Toury's Map, is in a position to formulate rules or norms which are prescriptive in nature. The descriptive norms generalized from corpus-based findings can be used to balance out the introspection-based prescriptive rules in the translation classroom. Laviosa (2008: 126) argues that "corpus design and compilation, as well as processing tools and procedural steps, are very similar to those employed in corpus-based descriptive research." In this way, learning is similar to doing research and translation teaching can be made less prescriptive with the help of corpora. This point will be examined in more detail in the following sections.

2. Corpus Linguistics: Corpus and Language Pedagogy

According to Biber et al. (1998: 236), the corpus-based approach can inform education in three ways, namely, the dissemination of findings from corpus-based studies, the development of educational materials, and the design of classroom activities. The corpus is often used as a resource for L2 writing and grammar teaching because of its strength in producing natural collocations together with contexts.

In the language teaching field, corpora have been used for pedagogical purposes for more than twenty years since Johns (1986, 1991) first advocated concordancing as a research tool for international students at the University of Birmingham to acquire vocabulary and grammar rules. Starting from the 1990s, his view of using corpora for language teaching is also echoed by Tribble and Jones (1990). At first, students were given only printouts of corpus concordance lists (Stevens, 1991), and later with the easy access to and maturity of computer technology, students were encouraged to consult corpora directly (Cobb 1997). Recent years have seen an increase in studies focusing on the application of corpora to language teaching (Granger et. al.

2002; Aston et. al. 2004; Sinclair 2004; Flowerdew 2009). These works cover a wide array of topics on the use of corpora for teaching purposes, including corpus-based language descriptions, corpus application in classroom settings, network-based language teaching, and learner corpora (cf. Sinclair 2004).

Unlike contrived language data derived from intuition, corpus studies employ authentic texts, or “language in use” (Sinclair 1991: 39), thus the findings are grounded on empirical description and analysis of solid data. For this reason, the corpus is regarded as a reliable source of representing actual language use because of its capacity for producing objective data (Biber 1993; Sinclair 1991). The introspection-based approach has been criticized by Sinclair, who states that the problem with “all kinds of introspection is that it does not give evidence about usage”(39). A corpus can clearly overcome such limitations.

Johns (1991, 1994) emphasizes the importance of using authentic texts to generate concordance lines in printouts so that students can discover and generalize the rules by themselves. The concept of DDL (Data-driven Learning) is also proposed by Johns, who argues that DDL as a learning process can “confront the learner as directly as possible with data” so that learners take on the role of a linguistic researcher where “every student is Sherlock Holmes” (Johns 2002: 108). One important underlying assumption of DDL is that students can improve general skills by using context to discover and deduce meaning (Hunston 2002: 170). Bernardini (2000), who works in both the fields of language and translation teaching, also points out the strengths of a corpus in that it provides rich sources of “serendipitous” learning activities, i.e., students can follow up any interesting observations that they happen across. DDL, inherently an inductive discovery learning process, is widely applauded by researchers (Barlow 1996; Sinclair 1997; Mindt 1997; Aston 1999).

To a certain extent, the usefulness of corpora for language teaching is now well-acknowledged as compared to two decades ago, when the corpus was rarely heard of in academia. Sinclair (2004: 2), for example, states that

“corpora, large and small, are seen by many teachers as useful tools and are being put to use more and more every day”. More specifically, the use of corpora in educational settings has two important functions.

For teacher-researchers, a corpus can be used as a tool to objectively evaluate students’ language competence by identifying and analyzing the frequently-occurring errors or deviant patterns of students’ language production (Granger 1994, 1998; Aston 1995). The research in this area is often centered on L2 students’ language productions, which is referred to in the field of applied linguistics as the study of interlanguage.^[3] At the same time, the corpus is also used by researchers to conduct investigations to compare textbooks with large corpora to see if they are representative of language in use. (Kennedy 1987a, 1987b; Mindt 1996, 1997)

For students and language learners, the value of a corpus is seen in its ability to present a large amount of authentic and natural data in KWIC (Key Word in Context) form. This has a special implication for L2 learners who are often not familiar with the target culture and language norms. L2 learners can also examine the corpus to confirm and disconfirm their own intuitions (Hunston 2002: 95). In this way, students can overcome the unnaturalness of contrived language.

3. Corpus and Translation Pedagogy

In the field of translation teaching, the potential of corpora for informing translation teaching is also recognized by researchers. As Hunston (2002: 123) observes, research into corpora and translation tends to focus on two areas: practical and theoretical. As has been discussed above, there is a unidirectional influence of descriptive and theoretical branches on the practical branch of translation studies. Traditionally, second language learning and translation are

treated as two inseparable activities since translation involves the mediation of two different languages. This is particularly the case when translating out of one's mother tongue. For this reason, Bernardini argues as follows:

Corpora have an important role to play in the education of translators, first as translation aides, as testified in the literature, secondly as sources of learning activities and of knowledge about the language, and thirdly and more importantly perhaps, as instruments through which approaches to language teaching and to translation teaching can be integrated into a coherent whole, with common aims and methods specific to this pedagogic setting

(Bernardini 2004: 97)

This indicates that translation, mostly a language-based activity, is closely related to one's good command of both source and target languages. In this respect, corpus is believed to be beneficial because it helps translators to understand and command the language. Or, as Aston puts it,

[B]y drawing attention to the different ways expressions are typically used and with what frequencies, corpora can make learners more sensitive to issues of phraseology, register and frequency, which are poorly documented by other tools.

(Aston 1999: 292)

To a certain extent, the aim of improving students' language competence (especially second language competence) is where corpus-based language and translation pedagogy intersect (cf. Bernardini 2003). In this respect, even a monolingual corpus can prove helpful.

In the field of applied translation studies, different types of corpora have been recommended and used for translation teaching purposes. A discussion of these different types of corpora is important for reviewing the

use of corpora in the area of translation teaching. In the following, a few types of corpora of immediate relevance to translation pedagogy are examined.

3.1 Monolingual Corpora

A monolingual corpus is one that contains texts in a single language. As translation involves two different languages, i.e., source language and target language, a monolingual corpus can be either one in the source language or one in the target language. This distinction is made by Coffey (2002) who refers to the former as SL (Source Language) monolingual corpus and the latter as TL (Target Language) monolingual corpus. Both corpora have been used by researchers for translation teaching purposes. On the whole, researchers welcome the application of monolingual corpora in translation teaching because they are generally larger and easier to construct. “A monolingual corpus is an equally valuable resource, though usually for different purposes. As monolingual corpora are generally larger and, in some cases, may be considered representative, they are able to offer information on more or less standard language use on the basis of quantitative data” (Vintar 2008: 153). In translation teaching practice, TL monolingual corpora are more popular among researchers since students are translating into the target language. Bowker (1998), for instance, uses a specialized monolingual corpus, and finds that corpus-aided translations are of higher quality in respect to subject field understanding, correct term choice and idiomatic expression than those using conventional resources such as dictionaries and encyclopedias. The use of TL monolingual corpora has been further explored by Bowker (2000), who demonstrates the advantages of corpora as a valuable resource for translators. By comparing two translations, one done using conventional resources and the other using a specialized corpus, Bowker concludes that “translators would be remiss not to use corpus resources” (17).

Stewart (2000), using a TL monolingual corpus of BNC^[4] (British National Corpus), studies how students translate tourist brochures from Italian

into English with the aid of this corpus. Stewart distinguishes between two types of translation: L1 translation (i.e., translation into the mother tongue) and L2 translation (translation into the foreign language), and hypothesizes that L2 translations tend to be more conservative than L1 translations. This is partly testified by his experience of using the BNC corpus as an aid to translation teaching, which tells him that students tend to use more frequently-occurring language patterns that the corpus provides. The issue of conventionality and creativity in corpus-assisted translation teaching is addressed, and Stewart cautions that:

1. It may be premature to introduce corpora in a course constituting the students' first exposure to professional, i.e., non-pedagogical translation training.
2. It seems counter-productive to introduce retrieval software into the courses if a committed training programme is not envisaged precisely for that purpose. Insufficient expertise in this area will infallibly result in clumsy, superfluous searches.

(Stewart 2000: 85-86)

The concept of conventionality and creativity is a controversial one. Some scholars are against corpus application in translation classrooms because a corpus, which is inherently conservative in nature, encourages the neutral use of language. Stewart, quoting Baker and Sinclair, while admitting that corpora are conservative, also argues that they can be used as a conventional backdrop against which creative usage can be measured (87). The training of students to master concordance software or make valid judgments as to what constitutes an appropriate translation choice is not mentioned in this respect. As Stewart mentions, "the searches made by the students were often gratuitous" and "students are too easily persuaded by recurring patterns highlighted in the corpus" (85). Since the corpus Stewart uses is

a general corpus instead of a specialized one and it contains only texts in one language, in this case, texts in the target language, the validity of his judgments is to be doubted.

Other researchers who have proposed the use of a TL monolingual corpus in translation teaching include Bowker and Pearson (2002) and Wilkinson (2005). However, their arguments are based on small-scale self-explored research instead of longitudinal empirical studies. In 2008, a group of professional translators in Spain (Maher et. al. 2008) argue that a corpus-guided approach, i.e., using a TL monolingual specialized corpus, can provide translators a faster and more economical way to master terminologies of a specialized field than other means. Their views, arising from actual translation practice, are worth noting in this respect.

The use of the SL monolingual corpus is advocated by Coffey (2002). He argues that a SL monolingual corpus can be helpful in two ways. First, on the part of translation students, it can function as a translation aid by providing linguistic and encyclopedic information. Second, on the part of translation teachers, such a corpus can be used as a source of teaching and testing materials. As a matter of fact, Coffey focuses on the use of a corpus from a teaching perspective rather than from a learning one, as he devotes a large part of his discussion to the ways in which the corpus can be exploited as a resource for teaching. The function of a SL corpus for providing linguistic and encyclopedic information is similar to that of a TL monolingual corpus or a comparable corpus. It is worth noting, though, that he points out that a SL monolingual corpus can be helpful to those translators whose mother tongue is not the SL.

In summary, most of the discussions on the use of monolingual corpora in translation studies are concerned with the potential of a corpus to provide linguistic information such as natural-sounding collocates and expressions. The use of such a corpus often presupposes competence on the part of translation students in comprehending the source language. Most

researchers use a monolingual corpus because parallel corpus is difficult to construct and rarely available. A monolingual corpus, in most cases, is chosen for the sake of convenience and easy construction. This is more or less also the case with a comparable corpus, which will be examined in the next section.

3.2 Comparable Corpora

In terminological terms, researchers differ in their understanding of the concept of comparable corpus. Comparable corpus is so named as it is based on two separable sets of texts which are thought to be comparable in nature. The two sets of texts, which can be of one language or two different languages, are two sub-corpora (c.f. Aston 1999: 291). There are at least three types of comparable corpora:

- Type A: One set of texts in language A and one set of texts in language B, which are comparable in terms of subject field or other attributes.
- Type B: Spontaneously-sourced texts (i.e. texts originally written) in language A and translated texts from other languages into language A.
- Type C: Two sets of different translated texts in language A, which originate from one single source text.

Generally, the comparable corpus used for translation teaching purpose is Type A (The comparable corpus discussed hereinafter refers to this type). Type B is often used by translation scholars as a means to identify translation universals. There is one particular reason why researchers choose comparable corpus as a resource for conducting translation teaching. In a sense, a comparable corpus falls in-between a monolingual corpus and a parallel corpus in terms of its design. A monolingual corpus clearly has its limitations since it can only provide references in one language, even though it might be of a considerable

size. On the other hand, a parallel corpus is difficult to construct and little commercially-available concordance software has been developed specifically for teaching purposes. Moreover, it is difficult to set up a sizable parallel corpus to significantly inform translation teaching. Indeed, the excessive labour and technical obstacles confronting parallel corpus construction have discouraged many from pursuing such an approach.

Most comparable corpora contain specialized texts, i.e., texts belonging to genres or domains which are socio-linguistically similar in each of the cultures involved (in terms of participation framework, function, and topic) (Aston 1999: 291). Because of its obvious advantages over monolingual corpora, a comparable corpus is introduced to the translation teaching field and its pedagogical value has been recognized by a number of researchers. In his article titled “Bilingual Comparable Corpora and The Training of Translators”, Zanettin (1998) argues that even a small bilingual comparable corpus made of either general or specialized language can be used to help students enhance their understanding of the source language text and produce fluent target language texts.

Generally speaking, a comparable corpus, which is made up of one SL monolingual corpus and one TL monolingual corpus, has all the advantages and attributes of monolingual corpora. Kübler (2003), combining different types of corpora and the Web for the training of terminologists, claims that the use of “comparable corpora in LSPs helps to overcome problems of ‘artificiality’ in parallel corpora”(41).^[5] Nevertheless, the parallel corpus mentioned here is of a small size (comprising Linux HowTo’s in English and translations in French) and it is far from being a satisfactory workable corpus due to its limited size. Kübler’s study, similar to a number of other studies, does not mention the important variables (size, representativeness, etc) which determine the extent of an effective parallel corpus. Like those who work solely with monolingual or comparable corpora, Kübler also makes subjective statements by praising the merits of a comparable corpus while dismissing the

usefulness of parallel ones.

By using a comparable corpus of English and Italian newspaper texts, Zanettin (2001) reports on a study in which a group of undergraduate students were asked to translate part of a newspaper article from Italian to English. He concludes that “[u]sing comparable corpora and concordancing software as aids in translation activities can help learners gain insights into the language and the cultures involved and develop their reading writing skills” (193-194). However, the study relies heavily on the author’s self-reports rather than actual evidence deriving from translation experiments, thus its reliability is questionable. Some important parameters of empirical research such as information of subjects, research methodology, and data analysis procedures are not included. This is the weakness of most corpus-related studies in educational settings. Other studies dealing wholly or partially with comparable corpora include Aston (1999), and Gavioli and Zanettin (1997), and Maia (2003). Nonetheless, most of these studies are seldom based on longitudinal systematic studies. Thus, their claims are not very forceful.

3.3 Parallel Corpora

As has been mentioned above, parallel corpora are less studied in comparison to other types of corpora for educational purpose. This is partly due to a lack of accessible bilingual concordancing software and a scarcity of parallel texts. The term “parallel corpus” is used here to designate a collection of texts in language A and of their correspondent translations into language B (cf. Baker 1995; Zanettin 2001:177-178). Often, a parallel corpus is aligned at the sentence level to allow a concurrent display of source text and its correspondent translations through keyword search. Although researchers mostly recognize the advantages of using a parallel corpus, they are also cognizant of the difficulties of constructing a workable parallel corpus. For example, as stated by Bowker (1998):

The merits of using bilingual and parallel corpora for translation purposes are obvious; however, the problem here is that there are relatively few corpora of this type available (certainly not enough to cover the wide variety of subject fields that translators have to deal with), and it is more difficult to create this type of corpus because it needs to be aligned. (633)

This understanding is more or less shared by researchers working in the field (c.f. Coffey 2002; Bowker & Pearson 2002; Wilkinson 2005). However, with the advancement of computer technology and increasing availability of bilingual texts, setting up a parallel corpus is not as technically problematic as before, though it still demands considerable effort. Considering the benefits that it brings, the construction of a sizable corpus is worth such effort. The reason why researchers opt for monolingual corpora is not only because it is easy to establish, but also because of the large amount and varieties of texts available for reference. Translated texts are sometimes criticized because they do not provide a full range of linguistic possibilities, and when they do, they are dismissed for being artificial and giving “a distorted picture of the language they represent” (Teubert 1996: 247). However, this judgment can be challenged since a parallel corpus nowadays can offer a large number of texts in both source language and target language. In this way, translators can form their own translation strategies by examining the varied number of choices generated by parallel corpus. Since a parallel corpus is based on the concept of equivalence, i.e., it is based on a bidirectional relationship between source texts and target texts, it can also be exploited by translators to examine how one translation is correspondent to a natural expression in the source language. The bidirectionality of parallel corpus means that translation students can use both the source texts and target texts as a reference to examine their own intuitions. In this sense, as far as the naturalness of language is concerned, half of the language data in a parallel corpus are natural-sounding, and in reality can be used as a TL reference corpus.

At the beginning, the exploitation of parallel corpora for pedagogical

use was mostly connected with terminology extraction and teaching (Danielsson & Ridings 2000; Pearson 2000a; Meyer et al. 2000). At the same time, it was recommended for teaching language because it could provide information in collocations, phrasal patterns and lexical polysemy (Barlow 2000), and for teaching translation because it can extract bilingual information of collocations and idioms (Peters et al. 2000). Pearson (2003), using a small parallel corpus of popular science articles translated from English into French, demonstrates through the translation of the names of universities that translation strategies differ, namely, some of these names were directly translated, some were not translated at all, while a few were translated by other means. The author argues that by examining the translation strategies of previous translators, students can use the corpus evidence to “draw up their own translation guidelines” (23).

In summary, most of these studies are based on a small-sized parallel corpus and the findings are often based on the researchers’ own intuition-based analysis and reasoning rather than on empirical studies or teaching experiments. To a certain extent, these studies are more conceptual than empirical in nature. No research has touched on the students’ own perceptions of corpus use in translation learning. This is mostly due to the scarcity of readily-available parallel corpus. Nevertheless, the merits of parallel corpora over comparable corpora are acknowledged by corpus researchers. This is aptly stated by Pearson (2003) as follows:

Thus, investigations of parallel corpora may allow students to see how writers, i.e., translators, behave when constrained by the existence of a text composed in another language. Translators have to act as cultural and linguistic mediators, negotiating their way between languages and between cultures. They have to gauge how much of the material in a source text is directly transferable to the target language, how much of it needs to be adapted or localized in some way, whether any of it can, or indeed should be omitted. The answers to questions

of this nature cannot be found in comparable corpora because these issues never arise in monolingual text-producing environment. They only arise because of the constraints of a text composed in another language. The answers must therefore be sought in parallel corpora. (17).

3.4 Ad-hoc or DIY Corpora

Other types of corpora that are related to translation teaching include the ad hoc corpora, also known as disposable corpora or DIY (do-it-yourself) corpora (cf. Bertaccini & Aston 2001; Varantola 2003; Zanettin 2001).

As suggested by its name, ad hoc corpora are corpora which are created ad hoc for a certain purpose. As the world is ever changing with new creations, translators are at times confronted with types of text they are not familiar with. This is when ad hoc corpora come into play. Ad hoc corpora exploit the World Wide Web for text materials to construct a topic-related raw corpus, which is often disposed after the translation project is completed.

Ad hoc corpora can fall into the categories of either monolingual or comparable corpora, but rarely parallel corpora. Varantola (2003) examines in detail the construction principles of ad hoc corpora by using internet resources. According to Varantola (2003: 56), “[ad hoc] corpora do not need to be sophisticated in terms of syntactic or semantic tagging. In fact, they can be structurally very simple, text-only corpora, but they can nevertheless be very useful in the actual decision-making process in translation”.

The value of ad hoc corpora was testified to by Bertaccini and Aston (2001) who use a self-compiled ad hoc corpus made up of 20 texts that contain the word *clochemerle*.^[6] The researchers then utilize this corpus to show how the meaning of the French word *clochemerlesques* in the source text is identified. It is demonstrated here that the use of ad hoc corpora can help translators to better understand the meaning of source texts.

Although the web is not considered by many researchers as a qualified corpus (Zanettin 2001; Sinclair 2005), it has long been exploited as a rich

resource for corpus materials (e.g. Pearson 1998; Varantola 2000, Bertaccini & Aston 2001; Maia 2002). However, it remains controversial whether the web should be a reliable source of providing materials as these are constantly changing and of dubious authority. Nonetheless, the techniques of selecting appropriate texts and drawing inferences are important for using the web as a corpus resource. (cf. Pearson 2000b)

3.5 Learner Corpora

Researchers also show interest in the development and use of learner corpora⁷¹ (cf. Olohan 2004: 172-3; Bowker & Bennison 2003). Learner corpora, also designated by Bowker and Bennison (2003) as Student Translation Tracking System (SSTS), are different from the other corpora discussed above. It is mainly used by translation researchers or trainers as a tool to systematically track the records of students' translation performance. Such corpora are "useful [and] interesting in translator training context" (103) as they can enable longitudinal studies to be "carried out to chart the progress of individual students or groups of students". The SSTS is also referred to by Bowker (2003) as CCBT (Corpus Created By Translators). CCBT "corresponds in many ways to the type of learner corpus used in foreign language learning", which is defined by Leech (1998: xiv) as "a corpus, or textual database, of the language produced by foreign language learners" (Bowker 2003: 169). For this reason, learner corpora are not a direct learning source for translation students, but a research apparatus for researchers, similar to the one employed to study translation universals as discussed in Section 2. The only difference between the two is that the texts studied in the learner corpus are students' translations instead of those produced by professional translators. In this respect, a learner corpus is deemed to be pedagogically valuable to teacher-researchers.

4. Issues and Debates in the Use of Corpus in Translation Teaching

Generally speaking, corpus use is seen as a welcome move in the translation teaching field by most researchers. As noted by Laviosa (2011: 24), “corpora are being widely and systematically used in translator training” in applied corpus-based translation studies. According to Bernardini et. al. (2003), corpus use can help learners develop awareness (realizing that there is a problem and formulating appropriate questions), resourcefulness (knowing where to look for solutions to a problem, and how to assemble resources), and reflectiveness (being able to interpret results, draw conclusions, etc.), which are regarded as what professional translators should possess (Bernardini 2004). Bernardini also cautions that “any approach to didactic use of corpora is bound to take corpus-user interaction into account” (10), while holding the belief that “corpus work should help future translators increase their autonomy and flexibility, and that such experience should prove educationally valuable as well as professionally advantageous” (9). Recognizing both the limitations and strengths of the corpus, Bernardini et. al. (2003: 11) sum up that “the greatest pedagogic value of the instrument lies, [...], in its *thought-provoking*, rather than *question-answering*, potential”.

The different types of corpora used in teaching translation have different impact; nonetheless, they are often mentioned under the umbrella term of corpus. This is noticed by Aston (1999), who comments on the shortcoming of monolingual and comparable corpora of not “generating hypotheses as to possible translations” (304). In this sense, due to the scarcity of parallel bilingual data, studies of parallel corpora for teaching translation are relatively few and an overwhelming number of studies are instead focused on the use of monolingual and comparable corpora.

One underlying issue of corpus use in translation teaching is whether it increases conventionality in language use. This debate, arising from language

learning and DTS research, has been addressed by a number of researchers. As a warning remark to the researchers applying corpora to translation studies, Tymoczko (1998: 658) states:

Researchers using CTS tools and methods must avoid the temptation to remain safe, exploiting corpora and powerful electronic capabilities merely to prove the obvious or give confirming quantification where none is really needed, in short, to engage in the type of exercise that after much expense of time and money ascertains what common sense knew anyway.

Similar arguments are put forth by researchers such as Stewart (2000) and Malmkjaer (2003). Stewart, in his own research of using comparable corpus for translation teaching, also broaches the topic of conventionality and creativity of corpus use. He argues that students are easily tempted to use recurring patterns they identify in the corpus, which means that corpus study encourages the use of conventional linguistic patterns. However, he also points out that corpus use, like dictionaries and thesauri, do not impose a choice upon students. Besides, a corpus can also be used as a backdrop where creativity can be measured (87). As Hunston (2002: 213) points out, corpora can be authoritarian or empowering. If they are thought to contain all that is possible in a language, their use may hamper creativity. What is important is to teach students discernment so that a corpus can be used to the effect that students can know under what circumstance they should obey a norm or breach it (cf. Bernardini et. al. 2003: 11). For this reason, Malmkjaer (2003) proposes a pseudo-subversive use of corpora in translator training. According to Malmkjaer:

It is not always obvious which corpus might help a translator solve a specific problem; corpus evidence might be misleading in some cases; and offering past linguistic behaviour as a model for the future flies in the face of the nature of

language and may, furthermore, stifle invention. it is worth exploring ways of using corpora which may seem subversive of standard uses. (119)

To a certain extent, Malmkjaer's approach is only valid when texts in the corpora are of poor quality or the language instances are rare or not found in the corpus. Such a claim is still to be tested and should not be based on a few examples. Nonetheless, Malmkjaer has reminded us on the potential pitfalls of corpus use.

In summary, the corpus is recognized by most as of use to translators and translation teaching in general. However, there are some different voices regarding the use of corpus for encouraging conventionality instead of creativity in translation. The issue of conventionality and creativity in language use will not be addressed here as it has been much discussed by many researchers stated above. As far as translation is concerned, whether translation should be rendered creatively or conventionally is also an issue that requires much deliberation, which I will not address here. As summarized by Bernardini et. al. (2003: 12): "Corpus resources and software tools are at most useful tricks in the translator's bag". In order to make good use of corpora, users need to master a number of necessary corpus-related skills.

5. Directions for future research

Recent research in corpus applications to translation pedagogy, as noted in this review, has led to some new insights into the implementation of this new translation teaching method. The contrasting perspectives and viewpoints on the application of corpora in translation teaching are explored and elaborated; however, some of the research is preliminary and gaps remain. In the following sections, I will focus my discussion on some research questions that can further our understanding of corpus-assisted translation teaching.

5.1 Pedagogically-oriented corpora

Although difficult to construct, the bilingual parallel corpus, compared to other types of corpora, is an ideal medium for implementing corpus-assisted translation teaching. This is mainly related to the nature of translation, which involves the mediation of two different languages. Notwithstanding the obvious advantages of parallel corpora, it is worth noting that most of the research at this stage, more or less, come under the constraints of the under-availability of sizable parallel corpora which are developed specially for pedagogical applications (cf. Bowker 1998). The few parallel corpora that are available are often developed for research instead of teaching purposes. This is the case with a number of studies stated above, which utilized research corpora for teaching purposes. While research yields positive results, this type of studies has underlying threats as students are forced to first equip themselves with time-consuming corpus skills before they can manipulate it for actual learning needs.

Given the linguistically dual nature of parallel corpora and the complexities of real teaching settings, more research is needed to address issues of parallel corpus compilation, as advocated by Pearson (2003). The parallel corpus is in fact the most under-researched area in translation teaching and little progress has been made in these years.

5.2 Empirical and longitudinal research

It can be seen from the above review that most claims made by corpus researchers are based on conceptual deliberation or personal experience rather than systematic, longitudinal-empirical studies. Though some studies make use of certain examples to support their findings and claims, these examples are often of a minor and preliminary nature. Although the arguments and explorations made by translation researchers are conducive to theorizing, empirical evidence grounded on carefully designed experimental research would be more persuasive. In view of the complicated and challenging

nature of teaching, it is hard to imagine how the corpus-assisted approach to translation teaching can develop into a full-fledged teaching method without the evidence-based findings from empirical research.

In many ways, the corpus is not an easy tool that can be quickly added to a translator's toolkit. The introduction of corpora might complicate or "problematise rather than simplify the task of (future) translators" (Bernardini et. al. 2003: 11). For this reason, researchers should foresee the challenges and difficulties that might arise during this new teaching and learning process. That is why longitudinal research is deemed important because it helps address issues and support methods in ways that are not possible with traditional cross-sectional approaches. For example, longitudinal research can look into attitudes and perceptions of corpus trainees from the learning instead of the teaching perspective. Translation students, as key players in the whole translation teaching setting, have seldom been studied and their roles are often ignored in prescriptive and theoretical models (Li 2002). In order to objectively evaluate the role of corpora in pedagogical settings, students' own assessment is important to verify whether a corpus-assisted approach is feasible and effective. Clearly, answers to these would require more empirical and longitudinal research.

5.3 Language Pairs

Empirical research of this type in non-European languages is yet to be conducted. Most of the existing research is based on European languages and the claims are yet to be verified in other language settings. In order to provide a full picture of corpus-assisted translation teaching, it will be worthwhile to conduct research to study how students in a different language context would react to corpus use in the translation classroom.

For example, Chinese is vastly different from European languages in many ways. Previous research has demonstrated the uniqueness of Chinese translated texts (e.g. Xiao, 2010). Naturally, the singularity of the Chinese

language more or less prescribes that corpus-assisted translation teaching would adopt a slightly different method. In many ways, more corroborative evidence derived from non-European languages is deemed important for establishing the corpus as a proper teaching medium in the translation classroom.

5.4 Educational theories

As has been noted at the beginning of this paper, corpus-assisted translation teaching comes under the influence of corpus linguistics, translation theories and educational theories. In the past, researchers often attached more importance to address issues related to corpus linguistics, such as searching corpus data for corpus compilation (Varantola 2003), using internet for extracting terminology (Kübler 2003) and compilation of archives of student translations (Bowker & Bennison 2003). Relatively speaking, researchers have seldom investigated the factors of translation theories and education theories in corpus application for pedagogical purposes. In particular, the role of educational theories was seldom examined as a variable in this type of research. The development of this research area should not underestimate the impact of educational research owing to the pedagogical nature of this new teaching method.

For example, from a teaching perspective, a number of variables can be thoroughly investigated to inform corpus application in translation teaching settings. These include trainees' L1 and L2 language proficiency level, task difficulty, mode of delivery, etc. More research into corpus-assisted translation teaching targeting different proficiency levels can also inform pedagogy. More in-depth studies that probe translators' decision-making processes, while engaged in different translation tasks, are important for the advancement of this research area. There is evidence that translating into and out of one's mother tongue are two vastly different mental activities (Campbell 1998). It might be interesting to investigate how and to what extent corpora can make an impact on translation quality when translation direction is introduced as a variable.

6. Concluding remarks

This article first made a thorough review of the state of the art of research on corpus application for translation research, language pedagogy and translation pedagogy. The different types of corpora as used for translation teaching purposes were introduced and their strengths and weaknesses were elaborated. Some issues and controversies surrounding corpus-assisted translation teaching were also addressed. Based on a thorough review of previous research in this area, some major areas for future research were proposed and discussed.

Developments in the research and understanding of the pedagogical application of corpora in translation teaching have accompanied general developments in translation studies, corpus-based translation research and developments in corpus technology. However, there is still room for more sophisticated and scientific investigations in order to fully justify corpus application in the translation classroom. It should be pointed out, though, that arguments for and against the use of corpora in translator training are of a theoretical nature at the current stage and lack empirical evidence. It still remains that empirical research in the field is sparse and does not adequately demonstrate the role of corpus in enhancing students' translation quality. For corpus-assisted translation teaching to be established as a solid teaching and research area, the conclusion of this review is that work is still needed to shore up the foundations. After all, greater recognition of this research area in the larger setting of translation studies and corpus linguistics can only be obtained through hard evidence grounded on systematic and empirical research.

Notes

- ^[1] The concept of non-translated texts is different from that of the source text. It refers to spontaneously sourced texts (Mason 2001: 72) in the target language. In other words,

the texts are written in the first place in the target language instead of translations from another language.

- [2] Often, a monolingual comparable corpus consists of two subcorpora, i.e., one set of translated texts and one set of non-translated texts, to enable comparability between the two.
- [3] The notion “interlanguage” was proposed by Selinker (1974). In his introduction to the term “interlanguage”, Selinker hypothesized “the existence of a separate linguistic system based on the observable output which results from a learner’s attempted production of a target language form” (1974:35).
- [4] The British National Corpus (BNC) is a 100 million-word collection of samples of written and spoken language from a wide range of sources, designed to represent a wide cross-section of current British English, both spoken and written. It can be accessed through URL <http://www.natcorp.ox.ac.uk/>.
- [5] The concept of parallel corpus is a controversial one because the term was sometimes used by some researchers to refer to a comparable corpus. In this case, it is used here to refer to the bilingual corpus which holds both original texts and their corresponding translated texts.
- [6] *Clochemerle* was a comic novel by G. Chevallier which ridicules factionalism in village politics, apparently well-known enough as an archetype of petty factionalism to be alluded to without explanation by French journalists.
- [7] In studies of Second Language Acquisition, a learner corpus is compiled and used to study the interlanguage produced by learners. In studies of translation teaching, the translations produced by student translators can also be compiled as a corpus to study their unique features and patterns. Hence, the concept of a learner corpus is used also by some translation researchers.

References

- Aston, G. (1995). “Corpora in Language Pedagogy: Matching Theory and Practice”. In *Principle and Practice in Applied Linguistics: Studies in Honour of H.G. Widdowson*. Eds. B. Seidlhofer, H.G. Widdowson and G. Cook. Oxford: Oxford University Press, 257-270.
- _____. (1999). “Corpus Use and Learning to Translate”. *Textus* 12: 289-314.

Corpora in Translation Teaching

- Aston, G., S. Bernardini, and D. Stewart (2004). *Corpora and Language Learners*. Amsterdam: John Benjamins.
- Baker, M. (1993). "Corpus Linguistics and Translation Studies: Implications and Applications". In *Text and Technology: In Honour of John Sinclair*. Eds. M. Baker, G. Francis, E. Tognini-Bonelli, and J. Sinclair. Amsterdam: John Benjamins, 233-250.
- (1995). "Corpora in Translation Studies: An Overview and Some Suggestions for Future Research". *Target* 7.2: 223-243.
- (1996). "Corpus-based Translation Studies: the Challenges that Lie Ahead". In *Terminology, LSP and Translation: Studies in Language Engineering in Honour of Juan C. Sager*. Eds. J.C. Sager and H.L. Somers. Amsterdam: John Benjamins, 175-186.
- Barlow, M. (1996). "Corpora for Theory and Practice". *International Journal of Corpus Linguistics* 1: 1-37.
- (2000). "Parallel Texts in Language Teaching". In *Multilingual Corpora in Teaching and Research*. Eds. S. Botley, T. Mcenery and A. Wilson. Atlanta: Rodopi, 106-115.
- Bernardini, S. (2000). "Systematising Serendipity: Proposals for Concordancing Large Corpora with Language Learners". In *Rethinking Language Pedagogy from a Corpus Perspective: Papers from the Third International Conference on Teaching and Language Corpora*. Eds. L. Burnard and A.M. Mcenery. Frankfurt am Main: Peter Lang, 225-234.
- (2003). "Designing a Corpus for Translation and Language Teaching: The CEXI Experience". *TESOL Quarterly* 37.3: 528-537.
- (2004). "Corpus-aided Language Pedagogy for Translator Education". In *Translation in Undergraduate Degree Programmes*. Ed. K. Malmkjær. Amsterdam: John Benjamins, 97-112.
- Bernardini, S., D. Stewart, and F. Zanettin (2003). "Corpora in Translator Education: An Introduction". In *Corpora in Translator Education*. Eds. F. Zanettin, S. Bernardini and D. Stewart. Amsterdam: John Benjamins, 1-13.
- Bertaccini, F., and G. Aston (2001). "Going to the Clochemerle: Exploring Cultural Connotations through Ad Hoc Corpora". In *Learning with Corpora*. Ed. G. Aston. Bologna: CLUEB, 198-219.
- Biber, D. (1993). "Representativeness in Corpus Design". *Literary and Linguistic Computing* 8.4: 243-257.
- Biber, D., S. Conrad, and R. Reppen (1998). *Corpus Linguistics: Investigating Language Structure and*

- Use. Cambridge: Cambridge University Press.
- Blum-Kulka, S., and E.A. Levenston (1983). "Universals of Lexical Simplification". In *Strategies in Interlanguage Communication*. Eds. G. Kasper and C. Faerch. London: Longman, 119-139.
- Bowker, L. (1998). "Using Specialized Monolingual Native-language Corpora as Translation Resource: A Pilot Study". *Meta* 43.4: 631-665.
- _____. (2000). "Towards a Methodology for Exploiting Specialized Target Language Corpora as Translation Resources". *International Journal of Corpus Linguistics* 5.1: 17-52.
- _____. (2003). "Terminology Tools for 'Translators'". In *Computers and Translation: a Translator's Guide*. Ed. H. L. Somers. Amsterdam: John Benjamins, 49-65.
- Bowker, L., and P. Bennison (2003). "Student Translation Archive and Student Translation Tracking System - Design, Development and Application". In *Corpora in Translator Education*. Eds. F. Zanettin, S. Bernardini and D. Stewart. Manchester: St. Jerome Publishing, 103-117.
- Bowker, L., and J. Pearson (2002). *Working with Specialized Language: A Practical Guide to Using Corpora*. London; New York: Routledge.
- Campbell, S. (1998). *Translation into the Second Language*. London: Longman.
- Cobb, T. (1997). "Is There Any Measurable Learning from Hands-on Concordancing?". *System* 25.3: 301-315.
- Coffey, S. (2002). "Using a Source Language Corpus in 'Translator Training'". in *TRALinea* 5 (Special Issue on 'Corpus Use and Learning to Translate'). Retrieved November 30, 2011 at <http://www.intralinea.it/intra/vol5/cult2k/coffey.htm>
- Dnaielsson, P., and D. Ridings (2000). "Corpus and Terminology: Software for the Translation Program at Göteborgs Universitet". In *Multilingual Corpora in Teaching and Research*. Eds. S. Botley, T. McEnery and A. Wilson. Amsterdam; Atlanta, GA: Rodopi, 65-72.
- Eskola, S. (2004). "Untypical Frequencies in Translated Language: A Corpus-based Study on a Literary Corpus of Translated and Non-translated Finnish". In *Translation Universals: Do They Exist?* Eds. A. Mauranen, and P. Kujamäki. Amsterdam: John Benjamins, 83-100.
- Flowerdew, L. (2009). "Applying Corpus Linguistics to Pedagogy: A Critical Evaluation". *International Journal of Corpus Linguistics* 14.3: 393-417.
- Frawley, W. (1984). *Translation: Literary, Linguistic, and Philosophical Perspectives*. Newark: University of Delaware Press; Associated University Presses.

Corpora in Translation Teaching

- Gavioli, L., and F. Zanettin (1997). "Comparable Corpora and Translation: A Pedagogic Perspective". Paper presented at The First International Conference on Corpus Use and Learning to Translate. Bertinoro, Nov, 1997.
- Granger, S. (1994). "The Learner Corpus: A Revolution in Applied Linguistics". *English Today* 39.10/3: 25-29.
- (1998). "The Computer Learner Corpus: A Versatile New Source of Data for SLA Research". In *Learner English on Computer*. Ed. S. Granger. New York: Longman, 3-18.
- Granger, S., J. Hung, and S. Petch-Tyson (2002). *Computer Learner Corpora, Second Language Acquisition and Foreign Language Teaching*. Amsterdam: John Benjamins.
- Hunston, S. (2002). *Corpora in Applied Linguistics*. Cambridge: Cambridge University Press.
- Johns, T. (1986). "Micro-concord: A Language Learner's Research Tool". *System* 14.2: 151-162.
- (1991). "Should You Be Persuaded: Two Samples of Data-driven Learning Materials". In *Classroom Concordancing (ELR Journal)*. Eds. T. Johns and P. King. University of Birmingham Centre for English Studies, 1-16.
- (1994). "From Printout to Handout: Grammar and Vocabulary Teaching in the Context of Data-driven Learning". In *Perspectives on Pedagogical Grammar*. Ed. T. Odlin. Cambridge: Cambridge University Press, 293-313.
- Johns, T. (2002). "Data-driven Learning: The Perpetual Challenge". In *Teaching and Learning by Doing Corpus Analysis: Proceedings of the Fourth International Conference on Teaching and Language Corpora, Graz, 19-24 July, 2000*. Eds. B. Kettemann and G. Marko. Amsterdam; New York: Rodopi, 107-117.
- Kennedy, G. D. (1987a). "Quantification and the Use of English: A Case Study of One Aspect of the Learner's Task". *Applied Linguistics* 8: 264-286.
- (1987b). "Expressing Temporal Frequency in Academic English". *TESOL Quarterly* 21: 69-86.
- Kenny, D. (2001). *Lexis and Creativity in Translation: Corpus-based Study*. Manchester: St. Jerome Publishing.
- Kübler, N. (2003). "Corpora and LSP Translation". In *Corpora in Translator Education*. Eds. F. Zanettin, S. Bernardini and D. Stewart. Manchester: St. Jerome Publishing, 25-42.
- Laviosa, S. (1998). "Core Patterns of Lexical Use in a Comparable Corpus of English Narrative

- Prose". *Meta* 43.4: 557-570.
- (2008). "Description in the Translation Classroom: Universals as a Case in Point". In *Beyond Descriptive Translation Studies: Investigations in Homage to Gideon Toury*. Eds. G. Toury, A. Pym, M. Shlesinger and D. Simeoni. Amsterdam: John Benjamins, 119-132.
- (2011). "Corpus-based Translation Studies: Where Does It Come From? Where Is It Going?". In *Corpus-based Translation Studies: Research and Applications*. Eds. A. Kruger, K. Wallmach and J. Munday. New York: Continuum, 6-27.
- Leech, G. (1998). "Learner Corpora: What They Are and What Can Be Done with Them". In *Learner English on Computer*. Ed. S. Granger. New York: Longman, xiv-xx.
- Li, D. (2002). "Translator Training: What Translation Students Have to Say". *Meta* 47.4: 513-531.
- Maher, A., S. Waller, and M. E. Kerans (2008). "Acquiring or Enhancing a Translation Specialism: the Monolingual Corpus-guided Approach". *The Journal of Specialised Translation* 10: 56-75.
- Maia, B. (2002). "Do-it-yourself, Disposable, Specialised Mini Corpora—Where Next? Reflections on Teaching Translation and Terminology through Corpora". *Cadernos de Tradução* IX: 221-235.
- (2003). "Some Languages Are More Equal Than Others: Training Translators in Terminology and Information Retrieval Using Comparable and Parallel Corpora". In *Corpora in Translator Education*. Eds. F. Zanettin, S. Bernardini and D. Stewart. Manchester: St. Jerome Publishing, 43-53.
- Malmkjær, K. (2003). "On a Pseudo-subversive Use of Corpora in Translator Training". In *Corpora in Translator Education*. Eds. F. Zanettin, S. Bernardini and D. Stewart. Manchester: St. Jerome Publishing, 119-134.
- Mason, I. (2001). "Translator Behavior and Language Usage: Some Constraints on Contrastive Studies". *Hermes* 26: 65-80.
- McEnery, T., and A. Wilson (2001). *Corpus Linguistics*. Edinburgh: Edinburgh University Press.
- Meyer, R., M. E. Okurowski, and T. Hand (2000). "Using Authentic Corpora and Language Tools for Adult-centred Learning". In *Multilingual Corpora in Teaching and Research*. Eds. S. Botley, T. Mcenery and A. Wilson. Amsterdam: Rodopi, 171-177.
- Mindt, D. (1996). "English Corpus Linguistics and the Foreign Language Teaching Syllabus". In *Using Corpora for Language Research: Studies in Honour of Geoffrey Leech*. Eds. J. Thomas

- and M. Short. New York: Longman, 232-247.
- _____ (1997). "Corpora and the Teaching of English in Germany". In *Teaching and Language Corpora*. Eds. A. Wichmann, S. Fligelstone, T. Mcenery and G. Knowles. New York: Longman, 40-50.
- Olohan, M. (2003). "How Frequent are the Contractions? A Study of Contracted Forms in the Translational English Corpus". *Target* 15.1: 59-89.
- _____ (2004). *Introducing Corpora in Translation Studies*. London; New York: Routledge.
- _____ and M. Baker (2000). "Reporting That in Translated English. Evidence for Subconscious Processes of Explicitation". *Across Languages and Cultures* 1.2: 141-158.
- Øverås, L. (1998). "In Search of the Third Code: An Investigation of Norms in Literary Translation". *Meta* 43.4: 571-588.
- Pearson, J. (1998). *Terms in Context*. Amsterdam: John Benjamins.
- _____ (2000a). "Teaching Terminology Using Electronic Resource". In *Multilingual Corpora in Teaching and Research*. Eds. S. Botley, T. Mcenery and A. Wilson. Amsterdam: Rodopi, 92-105.
- _____ (2000b). "Surfing the Internet: Teaching Students to Choose Their Texts Wisely". In *Rethinking Language Pedagogy from a Corpus Perspective: Papers from the Third International Conference on Teaching and Language Corpora*. Eds. L. Burnard and T. Mcenery. Frankfurt am Main; Berlin: Peter Lang, 235-239.
- _____ (2003). "Using Parallel Texts in the Translator Training Environment". In *Corpora in Translator Education*. Eds. F. Zanettin, S. Bernardini and D. Stewart. Manchester: St. Jerome Publishing, 15-24.
- Peters, C., E. Picchi, and L. Biagini (2000). "Parallel Corpora and Comparable Bilingual Corpora in Language Teaching and Learning". In *Multilingual Corpora in Teaching and Research*. Eds. S. Botley, T. Mcenery and A. Wilson. Amsterdam: Rodopi, 65-72.
- Selinker, L. (1974). "Interlanguage". In *Error Analysis: Perspectives on Second Language Acquisition*. Ed. J. C. Richards. Harlow: Longman, 31-54.
- Shlesinger, M. (1989). *Simultaneous Interpretation as a Factor in Effecting Shifts in the Position of Texts on the Oral-Literate Continuum* (unpublished MA thesis), Tel Aviv: Tel Aviv University.
- _____ (1991). "Interpreter Latitude vs. Due Process. Simultaneous and Consecutive Interpretation in Multilingual Trials". In *Empirical Research in Translation and Intercultural*

- Studies: Selected Papers of the TRANSIF Seminar, Savonlinna, 1988*. Ed. S. Tirkkonen-Condit. Tübingen: G. Narr, 147-155.
- Sinclair, J. (1991). *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.
- _____. (1997). "Corpus Evidence in Language Description". In *Teaching and Language Corpora*. Ed. A. Wichmann. New York: Longman, 27-39.
- _____. (2004). *Trust the Text: Language, Corpus and Discourse*. New York, N.Y.: Taylor and Francis.
- _____. (2005). "Corpus and Text - Basic Principles". In *Developing Linguistic Corpora: A Guide to Good Practice*. Ed. M. Wynne. Oxford: Oxbow Books, 1-16.
- Stevens, V. (1991). "Classroom Concordancing: Vocabulary Materials Derived from Relevant, Authentic Text". *English for Specific Purposes* 10: 35-46.
- Stewart, D. (2000). "Conventionality, Creativity, and Translated Text: The Implications of Electronic Corpora in Translation". In *Intercultural Faultlines: Research Models in Translation Studies*. Ed. M. Olohan. Manchester: St. Jerome Publishing, 73-91.
- Teubert, W. (1996). "Comparable or Parallel Corpora?". *International Journal of Lexicography* 9: 238-264.
- Tirkkonen-Condit, S. (2004). "Unique Items: Over- or Under-represented in Translated Language?". In *Translation Universals: Do They Exist?* Eds. A. Mauranen and P. Kujamäki. Amsterdam: John Benjamins, 177-186.
- Toury, G. (1995). *Descriptive Translation Studies and Beyond*. Amsterdam: John Benjamins.
- Tribble, C., and G. Jones (1990). *Concordances in the Classroom*. Harlow: Longman.
- Tymoczko, M. (1998). "Computerized Corpora and the Future of Translation Studies". *Meta* 43: 652-660.
- Vanderauwera, R. (1985). *Dutch Novels Translated into English: The Transformation of a "Minority" Literature*. Amsterdam: Rodopi.
- Varantola, K. (2000). "Translators and Disposable Corpora". In *Proceedings of CULT (Corpus Use and Learning to Translate)*. Bertinoro, Italy, November.
- _____. (2003). "Translators and Disposable Corpora". In *Corpora in Translator Education*. Eds. F. Zanettin, S. Bernardini and D. Stewart. Manchester: St. Jerome Publishing, 55-70.
- Vintar, Š. (2008). "Corpora in Translator Training and Practice: a Slovene Perspective". In *Incorporating Corpora: The Linguist and the Translator (Translating Europe)*. Eds. G.M. Anderman

- and M. Rogers. Clevedon: Multilingual Matters, 153-167.
- Wilkinson, M. (2005). "Using a Specialized Corpus to Improve Translation Quality". *Translation Journal* 9.3. Retrieved June 30, 2011 at <http://www.translationdirectory.com/article545.htm>
- Xiao, R. (2010). "How Different is Translated Chinese from Native Chinese? A Corpus-based Study of Translation Universals". *International Journal of Corpus Linguistics* 15.1: 5-35.
- Zanettin, F. (2001). "Swimming in Words: Corpora, Translation and Language Learning". In *Learning with Corpora*. Ed. G. Aston. Houston, TX: CLUEB, 177-197.

About the Author

Liu Kanglong is currently Assistant Professor in the Department of English Language and Literature, Hong Kong Shue Yan University. He received his PhD in Translation from The Chinese University of Hong Kong in 2009. He has published in a number of journals, including *Chinese Translators Journal*, *Foreign Languages and Translation*, *Journal of Translation Studies*, *Literary and Linguistic Computing*. He has practical and research experience in corpus-based translation studies, translation teaching, and empirical approaches to translation studies.