CORPUS-BASED TRANSLATION STUDIES: WHERE DOES IT COMES FROM? WHERE IS IT GOING?

Bozorova Vasila Ilhom qizi
Teacher at Uzbekistan state World Languages University

Abstract: *Baker was the first to propose using corpora to study translation and interpretation (1993). At the time, it was thought that translation theory would designate the fields of inquiry and develop operational hypotheses, while corpus linguistics would provide the mechanism for conducting empirical investigations in this new collaboration. The two partners would collaborate primarily for the purpose of advancing the discipline's descriptive branch. Since then, the collaboration has developed into a robust entity with an own denomination—corpusbased translation studies, or CTS. Its research encompasses a wide range of languages and study fields, from descriptive to applied studies. There have been previous discussions on the state of the art in corpus-based translation research (Laviosa 2002a). This essay (re)examines the relationship between corpus linguistics and CTS, as well as between CTS and descriptive translation studies (DTS), in light of current advancements. The goal is to determine which of the earlier assertions and forecasts are still valid and which areas of long-term CTS research show the greatest promise.*

Keywords: *corpara, corpus linguistics, mechanism, new collaboration, corpus-based translation*

INTODUCTION

Ten years after the inception of this quickly developing field of study, the first international conference dedicated solely to corpus-based translation studies (CTS) took place in July 2003 in Pretoria, South Africa. I had the chance to plan out CTS, evaluate its accomplishments objectively, and pinpoint potential lines of research for the future thanks to the event. The paper is divided into three sections, each of which is arranged chronologically and represents a significant period in the history of CTS. The field of translation studies saw its emergence in the first two years, 1993–1995; the years 1996–1999, in my opinion, marked the corpus linguistic turn; and starting in 2000, it is conceivable that CTS would take a turned toward cultural studies.

The dawn of corpus-based translation studies.

In 1993, Mona Baker contributed her groundbreaking work, "Corpus linguistics and translation studies: Implications and applications," to an anthology of research pieces honoring John Sinclair (Baker et al. 1993). According to Baker (1993), "translation scholars will be able to expose the nature of translated text as a mediated communicative event" due to the availability of vast corpora of both original and translated text and the development of a corpus-driven methodology." Two years later, the original concept was expanded upon by proposing particular research projects involving the design and analysis of parallel, bi/multilingual, and, most importantly, monolingual comparable corpora in Target's article "Corpora in translation studies: An overview and some suggestions for future research." It marked the beginning of a fresh alliance. It was the beginning of a new collaboration that would grow into a novel conceptual framework, just one of several that would invigorate

the field of translation studies in the 1990s. A plethora of complementary and conflicting theoretical ideas and methodologies emerged from the cross-fertilization of new disciplines of study such as pragmatics, critical linguistics, postcolonialism, gender studies, and globalization throughout this decade. Meanwhile, feminism, poststructuralism, polysystem, and skopos—all well-established fields of study—were thriving (Venuti 2000).

When the collaboration began, corpus linguistics—which Sinclair and Leech had described as "a new perspective on description" (Sinclair 1991, 2) and "a new way of thinking about language" (Leech 1992, 106)—was experiencing previously unheard-of growth. Super corpora of English, such the British National Corpus (BNC), Cambridge Language Survey, Longman Corpus Network, and Bank of English—the latter of which has more than 300 million words—were being assembled. These corpora possessed a minimum of 100 million words.

It was possible to provide visual and aural information simultaneously in each concordance line with the interactive Corpus of Spoken American English, for example (Chafe et al. 1991). New corpus types were also being created. Not to mention noteworthy was the fact that, by 1990, sixteen corpora totaling 365 million words were available, representing the languages of Europe. Additionally, major programs to create reference corpora for the Union's languages had begun to get support from the Council of Europe.

The methods and insights of corpus linguistics were influencing a wide range of applied linguistics fields, including lexicography, educational linguistics, machine and computer-assisted translation, contrastive analysis, terminology, forensic linguistics, and critical linguistics, to mention just a few. Given its remarkable track record of accomplishments, it seemed reasonable to assume that corpus linguistics would triumphantly enter the field of translation studies as well. Of fact, when Baker made her ideas in 1993 and 1995, the discipline was already familiar with corpora.

In reality, the first monolingual equivalent corpus of Swedish novels to explore translationese had already been assembled by Gellerstam at the University of Lund in 1986. Similarly, Lindquist conducted research on the Swedish translations of English adverbials using a parallel language database in 1989, also at Lund.

Their research found its place in practical studies since it aimed to use corpora as tools to enhance the translation process. Instead, a composite research program within the framework of descriptive translation studies (DTS) was being proposed in the early 1990s. DTS had advanced significantly, with the help of polysystem theorists like Even-Zohar, Toury, and Lefevere, to the point where Toury designated DTS as essential to the development of translation studies into a recognized empirical field in 1995.

"Without a proper descriptive branch, no empirical science can claim completeness and (relative) autonomy," according to Toury (1995, 1). One of the keys, if not the key, to the success of CTS, in my opinion, is the solid connections that were made over those years between corpus linguistics and DTS as a result of a number of shared problems arising from an empirical perspective. What concerns are these common ones?

Instead of focusing on idealized or intuitive language data, both disciplines of study examine real samples of language use; linguistic regularities are viewed as probabilistic

norms of behavior rather than prescriptive laws. Furthermore, because these linguistic patterns both reflect and replicate culture, they are intricately linked to sociocultural factors. Both corpus linguistics and DTS use a comparative research model wherein texts are analyzed across corpora representing various language varieties, such as translated versus non-translated language, original texts and their translations, different text types or different modalities within the same language, and so on. Descriptive hypotheses that assert the probabilistic generality of a given phenomenon are presented.

Therefore, the target text-oriented, historical-descriptive strategy created by Toury from polysystem theory and the corpus linguistic-descriptive approach proposed by Baker are both embraced by the same empirical paradigm.

Additionally, the corpus design principles and corpus linguistic analytical techniques were generally consistent with Toury's "discovery procedures," which entailed a slow, inductive transition from observable translational phenomena to the non-observable norms that influence translators' decisions. The approaches that Toury and Baker had envisioned for DTS and CTS were in the process of being developed and tested in the middle of the 1990s. Both of them emphasized the significance of creating a logical descriptive approach that would enable researchers to compare findings, repeat investigations, and methodically expand the current body of information regarding the nature of translation.

The corpus linguistic turn in translation studies.

At UMIST in Manchester, the first CTS analysis was performed in 1996. I developed a monolingual, multi-source English corpus as part of that project and provided a synthesis between a corpus-based technique and the simplification research, which was a line of inquiry that DTS followed in the 1980s. According to my research, there are basic lexical usage patterns in newspaper articles and narrative prose that are essentially unaffected by the source language and can be thought of as elements of translational English simplification. The use of tiny text samples and the absence of control for either text category or source language—two methodological flaws highlighted by earlier research—were corrected (Laviosa-Braithwaite 1996). Nearly simultaneously and immediately before other innovative syntheses were put out till the decade's end.

Toury's law of growing standardization states that when unique textual relations, like creative collocations, are developed in the source text, they are frequently substituted with conventional relations in the target text, like habitual or typical collocations, which causes the original set of textual relations to dissolve.

This is exactly what Scott (1998) discovered in the translation of A hora da estrela, where the translator's decisions—whether intentional or not, required or not—caused the author's peculiar repetition of the word nao to break up its cumulative effect, fading the sense of void evoked in the original text. When Sinclair's (1991) and Toury's (1995) ideas are combined, the rule of rising standardization will predict that the open-choice target text's default idiom principle frequently takes the place of the original text's principle.

Once more, this is not an absolute rule; rather, its application depends on a number of variables, including age, the degree of bilingualism, the translator's training and experience, and the status of translation in the target culture. For instance, the more peripheral a

translation is in the target culture, the more it will conform to established models and repertoires in the language.

In light of Kenny's (1999) findings, one might speculate that, similar to how the idiom principle functions in original texts, text type plays a significant role in the functioning of the law of rising standardization in translation.

In actuality, her corpus was intended to contain just modern experimental fiction; an alternative design might have produced a of imaginative pairings.

REFERENCES:

1. Aluisio, S. M., G. M. Pinheiro, M. Finger, M. Das Gragas, V. Nunes and S. E. O. Tagnin. 2003. The Brazilian Portuguese Corpus Lacio-Web: Overview and issues in corpuscreation. In Proceedings of the Corpus Linguistics 2003 UCREL, Technical Papers,ed. D. Archer, P. Rayson, A. Wilson and T. McEnery, 16 (1):14-21.

2. Aubert, F. H. and S. E. O. Tagnin. 2003. A corpus of sworn translations for linguistic and historical research. In Proceedings of the Corpus Linguistics 2003 UCREL,Technical Papers, ed. D. Archer, P. Rayson, A. Wilson and T McEnery, 16(1):54-61.

3. Baker, M. 1993. Corpus linguistics and translation studies: Implications and applications.In Text and technology: In honour of John Sinclair, ed. M. Baker, G. Francis and E.Tognini-Bonelli, 233-250, Amsterdam: John Benjamins.

4. Baker, M.1995. Corpora in translation studies: An overview and some suggestions for future research. Target 7 (2):223-243.

5. Baker.M. 1996a. Corpus-based translation studies: The challenges that lie ahead. In Terminology, LSP and translation: Studies in language engineering in honour of Juan C. Sager, ed. H. Somers, 175-186, Amsterdam: John Benjamins.

6. Baker.M. 1996b. Linguistics and cultural studies: Complementary or competing paradigms in translation studies? In Übersetzungswissenschaft im Umbruch: Festschrift fur Wolfram Wilss, ed. A. Lauer, H. Gerzymisch-Arbogast, J. Haller and E. Steiner, 9-19, Tübingen: Gunther Narr.

7. Baker.M. 1998. Réexplorer la Langue de la Traduction: Une Approche par Corpus. Meta 43(4):480-485. (Special issue. The corpus-based approach ed. S. Laviosa.)

8. Baker.M. 1999. The role of corpora in investigating the linguistic behaviour of professional translators. International Journal of Corpus Linguistics 4 (2):281-298.

9. Baker.M. 2000. Towards a methodology for investigating the style of a literary translator. Target12 (2):241-266.

10. Baker.M. 2002. Corpus-based studies within the larger context of translation studies, Genesis:Revista cientifica do ISAI 2:7-16.

11. Bassnett, S. 2001. The translation turn in cultural studies. In Athanor XII (4):151-164. (Special issue: Lo Stesso Altro, ed. S. Petrilli.)

12. Bassnett, S. and A. Lefevere. 1998. Constructing cultures. Essays on literary translation. Clevedon: Multilingual Matters