Corpora: A Lens into Translation Phenomena

Sara Laviosa[1] Kanglong Liu[2]

Address: [1]Department of Humanistic Research and Innovation, University of Bari 'Aldo Moro', Bari, Italy;
[2]Department of Chinese and Bilingual Studies, The Hong Kong Polytechnic University, Hong Kong, China
E-mail: [1]sara.laviosa@uniba.it; [2]kl.liu@polyu.edu.hk
Correspondence: Sara Laviosa

In the late 1990s, when the corpus-based approach to the study of the product and process of translation and interpreting was still pioneering work, Maria Tymoczko (1998) predicted that it would play a leading role in the discipline of Translation Studies in the following decades. One of the main reasons for imagining such scenario was the capacity of corpus studies to "change in a qualitative and quantitative way both the content and the methods of the discipline of Translation Studies, in a way that fits with the modes of the information age" (Tymoczko 1998, 652). Hence, as Tymoczko contended, "Corpus translation studies is central to the way that Translation Studies as a discipline will remain vital and move forward" (Tymoczko 1998, 652).

Indeed, Tymoczko's predictions have all come true. Since its advent, Corpus-based Translation Studies (henceforth CBTS) has amply demonstrated that one of its strengths is the capacity of continually adapting modern technologies to the discipline's ever-changing needs and purposes in line with socio-cultural changes and the demands of the job market. What are the emergent exigencies of Translation Studies that CBTS sets out to address today? What are the new goals of Translation Studies that CBTS intends to pursue? The papers commissioned for this Special Issue of *Translation Quarterly* aptly illustrate how CBTS is exploring new avenues of enquiry in order to meet the current demands of Translation Studies and contribute to the achievement of its current objectives. CBTS is capable of doing so thanks to the design and availability of new corpora in different languages as well as the use of sophisticated methods of statistical analysis.

The papers we have selected for this Special Issue of the journal constitute a representative sample of recent trends in CBTS, that arise from the needs of our globalized world and

an increasingly technologized language industry worldwide. These trends pertain to two main research domains, namely descriptive and applied studies. Within descriptive CBTS, we can identify the quest for translation universals as a line of enquiry that builds on the theoretical and empirical work of previous scholars, continually refines the methodology used, and discovers patterns of translational language that enhance our understanding of the nature of translation as a distinctive form of cross-linguistic and cross-cultural mediation that is influenced by the stylistic norms established for particular genres, registers or language varieties. Indeed, as Kirsten Malmkjær (2008) suggests, general features of translational behaviour such as simplification, explicitation and normalization would be better accounted for by the norm concept and explained on socio-cultural grounds.

Within applied CBTS, research into machine translation, as part of translation aids, goes hand in hand with research into translation quality assessment, as part of translator training. Translator education in turn comprises the subdomains of teaching methods, testing techniques and curriculum planning. These ambits of scholarly enquiry aim to respond to the challenges posed by the growing impact of technology in the provision of translation services, particularly in the last decade. Another research endeavour that is making inroads into applied translation studies, especially in translation pedagogy, arises from the need to develop the ability to translate into language B. This concern is well motivated by the continuing expansion of English as lingua franca.

The first paper by **Sara Laviosa and Kanglong Liu**, "The Pervasiveness of Corpora in Translation Studies", provides the general background to the four contributions that follow on. The first author gives an overview of corpus use in Translation Studies from the early 1990s to the first decade of the new millennium. The second author assesses the state of the art of corpus studies of translation over the last ten years. In the concluding section, they pull together the main threads they laid out in their overview of the field, and then point to future directions.

In the second paper, "Uncovering Machine Translationese Using Corpus Analysis Techniques to Distinguish between Original and Machine-Translated French", **De Clercq et al.** analyse the linguistic features that distinguish original French newspaper texts from English to French machine-translated newspaper texts, using four machine translation systems, that is one statistical and three neural systems. The aim is to evaluate the output of different types of MT systems for both translator training and professional purposes.

In the third paper, "A Corpus-based Approach to Profiling Translation Quality: Measuring and Visualizing Acceptability of Student Translation", **Yanmeng Liu** addresses the thorny issue of assessing the quality of student translations in a valid and reliable way. To achieve this goal, Liu develops a corpus-based approach to profiling and assessing translation acceptability, conceived as a dimension of translation quality. The approach involves the statistical comparison of lexical, syntactic and grammatical features of English texts translated from Chinese vis-à-vis original English texts. More specifically, Liu uses two corpora, i.e. the

Parallel Corpus of Chinese EFL Learners (PACCEL) and the Lancaster-Oslo/Bergen Corpus (LOB), together with machine learning methods and analytical statistics to verify assessment efficiency and display the results.

In the fourth paper, "Translation Universals in Legal Translation: A Corpus-based Study of Explicitation and Simplification", **Francesca Luisa Seracini** combines quantitative and qualitative analyses of patterns of language use in translated EU texts in Italian versus original English EU texts. These typical features of translational language are considered to be manifestations of the posited translation universals of simplification and explicitation. Seracini argues that the occurrence of these linguistic patterns provides not only evidence for the presumed existence of distinctive features of translational behaviour regardless of the influence of the source language, but also evidence for the influence of the stylistic norms recommended for given subject-specific domains in the target language and culture, such as legal language.

In the fifth and last paper, "English for a Global Readership: Implications for the L2 Translation Classroom", **Dominic Stewart** tackles the complexity of establishing what are appropriate or inappropriate translation renderings when the target language is English as a Lingua Franca (ELF). Based on examples drawn from his Italian-to-English translation classes, where different corpus-based resources are regularly used for translating tourist text aimed at an international readership, Stewart demonstrates the conundrum that teachers and students face when different reference materials (such as monolingual dictionaries for advanced learners of English) and different corpus-based resources (such as English reference and web-derived corpora) provide inconsistent evidence as to the accuracy and fluency of potential translation equivalents.

This special thematic issue "Recent Trends in Corpus-based Translation Studies" fittingly closes with a review by Cui Xu of an edited volume, *CTS Spring-cleaning: A Critical Reflection*, edited by María Calzada Pérez and Sara Laviosa. Published in a Special Issue of the journal *MonTI* in 2021, the articles indeed highlight the recent trends in the field of CBTS by placing a special focus on overlooked areas such as subtitling, travel journalism, localization, and operatic audio description. As stated by Xu in the book review, the studies have offered refreshing perspectives in the field of corpus-based translation and interpreting studies.

To conclude, based on the evidence provided by the papers collected here, we can confidently affirm that CBTS is proving to be exactly what Tymoczko (1998, 652) envisaged it would be. CBTS is promoting "the construction of information fields that suit a new international, multicultural intellectualism, providing for the inclusion of data from small and large populations, from minority as well as majority languages and cultures". CBTS is stimulating collaborative research endeavours "unimpeded by time or space". Like large databases in science, corpora make it possible to build upon past research and become "a legacy of the present to the future, enabling future research to build upon that of the present". CBTS has marked "a turn away from prescriptive approaches to translation toward descriptive approaches". Fi-

nally, CBTS is reengaging "the theoretical and pragmatic branches of Translation Studies, branches which over and over again tend to disassociate, developing slippage and even gulfs" (Tymoczko 1998, 658). Indeed, the technical and descriptive investigations undertaken within corpus-based studies have now, more than they have ever had, "practical potential and immediate applicability, not only for the teaching of translation but for the work of the practising translator as well" (Tymoczko 1998, 658).

References

Malmkjær, Kirsten. 2008. "Norms and Nature in Translation Studies." In *Incorporating Corpora. The Linguist and the Translator*, ed. by Gunilla Anderman, and Margaret Rogers, 49-59. Clevedon: Multilingual Matters.

Tymoczko, Maria. 1998. "Computerized Corpora and the Future of Translation Studies." *Meta: Journal des Traducteurs/Meta: Translators' Journal* 43(4): 652-659.