REVIEW

Sylviane Granger and Marie-Aude Lefer: *Extending the Scope of Corpus-based Translation Studies*

(London: Bloomsbury, 2022. 274 pages)

Jiaqi Tan

Xi'an International Studies University
China
iristjq@126.com

Rongbo Fu

Ningbo University
China
frbjoe@126.com

68
June
2025

Empirical translation studies in the past few decades have benefited profoundly from the advances of computer technologies. Among them is the wide use of electronic machine-readable corpora, which provides major impetus for the development of Corpus-based Translation Studies (CBTS) pioneered by Mona Baker's seminal call three decades ago. While research on translation has gained much more fresh insights from the compilation and application of a wealth of parallel and monolingual corpora involving different languages, there seems an increasing need for theoretical and methodological rigor in conducting CBTS, aiming to help address the analysis of various new and diverse corpus data. In this regard, the present collection of articles under review is certainly a much-awaited contribution to this area.

The volume, consisting of 9 chapters structured in four themed sections, offers a comprehensive overview of current and future trends in CBTS by taking stock of innovative attempts in terms of theoretical, methodological, empirical and pedagogical explorations.

Chapter 1 starts with Granger and Lefer's retrospection of the history of CBTS. Based on a selection of 186 evidence-based articles from twelve scientific journals over an eight-year period, their bibliometric analysis underlines a methodological weakness in existing research, that is, an inadequate usage of more advanced corpus techniques and multivariate statistics. More importantly, the readers are constantly reminded of the necessity of manual filtering as a crucial step, albeit laborious and time-consuming, to exclude studies claiming to fall within the domain of corpus-based translation and interpreting studies yet fail to use electronic data and corpus analytic methods.

Chapter 2 by Gaspari's aims to suggest ways through which future CBTS can be expanded so as to address the lacunae and challenges facing researchers as a result of increasing availability and diversity of natural translation data. As an initial step, he focuses on the issues of translation (and mediation) universals (e.g., explicitation) and directionality which have attracted much attention in CBTS over the last three decades. This is followed by his appeal that new methods need to be developed to keep up with the emergence of new forms of translation (e.g., amateur translation, translation crowdsourcing, game localization, trainee and expert interpreters' training).

Chapters 3 and 4 are dedicated to exemplifying fresh theoretical and methodological approaches in CBTS and each serves as a case study of how they can be meaningfully applied to enrich CBTS. Chapter 3, contributed by Kotze, maps out three core principles and five constraint dimensions to reinforce the theoretical basis for the constrained language framework and reports on a replication study as a way to highlight the framework's potential for translation studies. Specifically, Kotze revisits the phenomenon of the optional *that* complementizer by making a corpus-based, multifactorial analysis of the factors conditioning the choice between *that*/zero in three English varieties (English translated from Afrikaans, Brit-

ONOMÁZEIN 68 (June 2025): 224 - 228
Jiaqi Tan and Rongbo Fu
Sylviane Granger and Marie-Aude Lefer: *Extending the Scope of Corpus-based Translation Studies*

226

ish English and South African English). Her efforts in developing the constrained-language framework with increased breadth and depth provide further insight into the interplay between various kinds of constraints and various forms of language production, which has profound implications for multiple constraints behind translators' decision-making process particularly compared with other different varieties of constrained language. In Chapter 4, Neumann *et al*. focus on the use of multi-method design in CBTS, that is, combining corpus, experimental and behavioral data in the examination of cross-linguistic features in translational products. Following an in-depth discussion on integrating data as such, they apply the methodology to the study on the role of subject identifiability and thematic status on positional shifts in English-to-German translation. It turns out that the results of corpus are only partially consistent with those of the experiment, suggesting that different types of data may offer complementary insights into the process of translation. Here, the authors' elaboration on how observational and experimental data combined can be seen as a useful guide for conducting research as such, which was less sufficiently discussed in previous studies calling for convergence of evidence from process- and product-oriented approaches such as Liang (2020) and Kotze (2020).

Chapters 5, 6 and 7 report on three fresh empirical endeavors using cutting-edge data-mining tools and analysis methods to expose semantic and syntactic particularities of the translated language versus non-translated language. Chapter 5 by Ivaska *et al*. takes, under the constrain-language framework, a corpus-driven approach to syntactic properties of constrained English. By availing themselves of a purpose-built dataset of three English varieties (English as second language, translated English and non-translated English as a first language) featuring three registers (argumentative writing, political speeches and tourism-related communication) and two first/source languages (German and Italian), the authors aim to examine the degree to which these constrained languages differ in terms of part-of-speech (POS) dependency bigrams, a more advanced measure of syntactic structure rarely used in translation studies. Their study lends weight to the hypothesized interplay between constrainedness and impact of registers: cross-register differences are less noticeable in constrained varieties than unconstrained ones. In Chapter 6, Heilmann *et al*. investigate how various source-language variables condition the alternation between nominal re-metaphorizations and de-metaphorization of *of*-construction based on a bi-register English-to-German parallel corpus. Their findings suggest that only the semantic category of ENGAGEMENT significantly affect translation shifts from a non-congruent variant to a congruent version, hinting at the fact that keeping the level of metaphoricity may be a norm in translation from English to German. Chapter 7 by Lapshinova-Koltunski probes into the influence of translation experience and register variation on the effects of normalization and shining through. To do so, she builds a multi-register parallel corpus comprising English originals and their German translations by students and professionals. Drawing upon theoretical insights from Systemic Functional Linguistics, she then operationalizes the two linguistic phenomena into a set of lexico-grammatical

features whose extraction could be facilitated using the Corpus Query Processor. Her conclusion is that the translationese effects tend to vary more significantly across different registers than across levels of translation expertise.

The last two chapters (8 and 9) delves into the potential usefulness of corpus in translator training, especially on judging the quality of student translators' performance. In Chapter 8, Verplaetse compares, using key-logging data, the benefits of a monolingual target language corpus (MOC) with a bilingual concordancer (Linguee) in translation classroom. She hypothesizes that the use of MOC would generate fewer acceptability errors and more adequacy errors than Linguee in students' translations. However, while findings confirm the positive impact of Linguee on enhancing target text adequacy, results of error analysis contradict the presumed benefit of the MOC in lowering acceptability errors versus the Lingee condition, suggesting the usefulness of monolingual resources against bilingual resources in improving students' translation quality is far from certain. Chapter 9 by Kübler *et al.* is also an experimental study that compares the use of corpus resources verses bilingual dictionaries in facilitating the students in their translation of specialised texts, focusing on the specific difficulty of rendering English complex noun phrases (CNPs) into French. The results seem unexpected: there is no significant difference in translating complex CNPs between the two conditions, namely with and without corpus, contradicting earlier findings which favor the efficiency of corpus use in trainees' translation practice. This leads the authors to design some remedial classroom activities which intend to help the students overcome the difficulties associated with the translation of complex CNPs even with the assistance of corpora and/or machine translation.

Taken together, contributions in this volume best epitomize the latest attempts at pushing boundaries of CBTS on various levels, ranging from conceptual to methodological and to pedagogical. They also testify the limitless potential of CBTS as a paradigm, which is able to continue evolving as a result of its interaction with other neighboring fields beyond corpus linguistics, such as psycholinguistics, cognitive linguistics, contrastive linguistics and variational linguistics, second-language acquisition, both in terms of research methods (e.g., mixed-methods design and multifactorial analysis) and theoretical insights (e.g., the constrained language framework and cross-linguistic contrasts in translation).

Apart from the theoretical and methodological advancements concerning CBTS it demonstrates, the edited volume has three major features to be specially highlighted. First, the inclusion of an impressive array of new research data including eye-tracking and keystroke logging data, post-editing data, student translations, which are less involved in earlier corpus-based work as such yet which are shown to be desirable as their combination increases "the validity, strength, and interpretative potential of a study" (Thurmond, 2001: 253). Second, the account of semantic, and syntactic aspects of translated language such as the semantic categories of the nominal *of*-constructions and POS dependency bigrams which have been

rarely explored in traditional studies of translated language. Third, the experimental application of corpora to translation training and quality assessment, which calls into question the efficiency of corpus resources on the one hand and sheds new light on the sources of errors influencing students' translational choices. All these make the volume under review a worthy reference for anyone who is interested in the use of corpus methodologies in the study, practice and pedagogy of translation.

That said, the volume would have been benefited from the addition of one or two studies on the use of interpreting corpora, an under-represented area of translation studies requiring separate coverage (Shlesinger, 1998). This, however, should be only a matter of time if hurdles regarding data collection, transcription and annotation, etc. can be gradually removed.

References

Kotze, Haidee, 2020: "Converging What and How to Find Out Why: An Outlook on Empirical Translation Studies" in Lore Vandevoorde, Joke Daems and Bart Defrancq (eds.): *New Empirical Perspectives on Translation and Interpreting*, Abingdon: Routledge, 333–371.

Liang, Junying, and Qianxi LV, 2020: "Converging What and How to Find Out Why: An Outlook on Empirical Translation Studies" in Lore Vandevoorde, Joke Daems and Bart Defrancq (eds.): *New Empirical Perspectives on Translation and Interpreting*, Abingdon: Routledge, 303–332.

Shlesinger, Miriam, 1998: "Corpus-based interpreting studies as an offshoot of corpus-based translation studies", *Meta* 43(4): 486-493.

Thurmond, Veronica A., 2001: "The Point of Triangulation", *Journal of Nursing Scholarship* 33 (3): 253–258.