Digital Humanities Approaches to Literary Translation

Yifeng Sun, Dechao Li

DIGITAL HUMANITIES APPROACHES TO LITERARY TRANSLATION

Yifeng Sun and Dechao Li

ABSTRACT

Digital humanities (DH), which first started as the compilation, digitization, and visualization of ample amounts of archival works of particular writers and artists (such as the projects of Rossetti Archive and William Blake Archive) at the beginning of the twenty-first century, has witnessed exponential expansion over the last decade. The recent rapid advances in modern computing and information technology have given rise to more new digital tools and methods, including hypermedia, data visualization, text mining and digital mapping, and so on, all of which have been applied in DH to foster systematic and scholarly study of these digital resources. Such a study has a distinctive interdisciplinary nature.

KEYWORDS: digital humanities, literary translation, corpus-based approaches, computing technology, translator's style

In addition to its earlier preoccupation with the building of "linguistic corpora, interactive digital archives and editing projects"[1] especially of historical and cultural documents, digital humanities (DH) today has demonstrated a much wider application in such traditional humanities disciplines as philosophy, linguistics, literature, and cultural studies, which used to be susceptible to individual researchers' interpretation based primarily on close reading, questioning, and analysis of a small number of nondigital materials (usually printed texts). Influenced by the macroscopic trends or patterns as revealed by the unprecedented amount of digital data that has usually been tagged, annotated, or aligned beforehand, scholars adopting the DH perspective

in these areas are now able to conduct humanistic inquiries into a broader range of interesting topics that can be fruitfully examined only through both quantitative and qualitative means.

Ning Wang[2] and Qinglong Peng[3] in this special issue point out how the field of literary research, including comparative literature studies, has been greatly innovated and expanded by the use of tools and methods of DH such as conducting algorithmic analysis of themes and text mining in literary texts. In fact, the impact of DH on literary studies has become so far-reaching that the DH methodology has even been applied in literature pedagogy to engage students in research and learning of various literary topics in this digital age.[4]

With DH gaining increasingly more momentum in comparative literary studies, its technology and research methodology have progressively reshaped literary translation. Admittedly, the study of literary texts in their original languages is generally encouraged or even preferred in the field of comparative literature, but it is also a fact that a fairly large number of students and scholars, especially those who are non-English speakers, still reply heavily on translated texts for their scholarly enquiry of literary works written in English. Thus, a comprehensive review of how DH has transformed the practice and scholarship of literary translation and also a synoptic outline of how it might continue to shed new light on this subfield of translation studies seem quite necessary, especially if we want to gain a more in-depth understanding of the relationship between literature and other social and cultural activities across linguistic and national boundaries.

The Early Attempts of Using DH in Literary Translation: Corpus–Based Approaches to Literary Translators' Styles (the 1990s *to the* 2000s*)*

Tracing the history of humanities computing—the former name of DH today, Susan Hockey summarizes its development into five phrases: beginnings (1949 to early 1970s), consolidation (1970s to mid-1980s), new developments (mid-1980s to early 1990s), and the era of the Internet (early 1990s onwards).[5] Although the first use of computer technology in the analysis of humanities works dates back to as early as 1949, when an Italian Jesuit priest Roberto Busa produced a computer-generated concordance of the writings of Thomas Aquinas by collaborating with IBM,[6] which represented the

first systematic attempt to apply computing technology to analyze literary translation at the dawn of the Internet age.

Before the 1990s, studies of literary translation, or even translation as a whole, were frequently susceptible to researchers' subjective or impressionistic interpretations that were based on limited and isolated examples. While such personal scrutiny might from time to time yield some interesting findings, it frequently ignores the regularities of behavior of translators as well as the general patterns of translational language. This speculative, free-standing approach to the study of translation is under severe attack by Gideon Toury, who instead calls for the development of a descriptive branch of translation studies (DTS) that aims at describing and explaining "observable or at least reconstructable facts of real life."[7] By "facts," he means the translation works that have found their ways into the target culture.[8] This empirical approach as advocated in DTS as well as the rapid developments of corpus linguistics at the beginning of the 1990s inspired some scholars to look to the potential of digital technology in the form of electronic corpora in the study of translation, including literary translation.

These include early attempts to investigate the so-called translation universals that are believed to be "general, all-encompassing tendencies"[9] in the use of language by translators irrespective of their nations, cultures, and languages, in translational corpus of literary texts. Mona Baker, who pioneered what was later known as corpus-based translation studies (CTS),[10] was also the first scholar to use corpus-based methodology to examine whether such linguistic tendencies as explicitation, simplification, normalization, and so on, also existed in translations from other languages into English, including literary translations. However, it has to be pointed out that although literary translations had been widely studied for general linguistic tendencies within the new paradigm of CTS since the 1990s, the research findings were most often used to support the general applicability of the hypotheses of translation universals for all translational genres, rather than to just illustrate the special features of translated literary language.

The fact that the language of translated literature has not become a legitimate object of study in its own right in CTS was not much changed until 2000, when Mona Baker proposed the use of digital methods to investigate translator's style in literary translation. In her paper published in that year, she proposed a corpus-based methodology for investigating a literary translator's style, which she meant the "thumb-print that is expressed in a range of linguistic—as well as non-linguistic—features"[11] of the translator. By comparing the standardized type-token ratio (STTR) and the mean sentence length as shown in the two subcorpora of English translations of

literary works by Peter Bush and Peter Clark, respectively, Baker found a significant difference in the corpus statistics of these two authors. In Bush's translations, he consistently registered a much higher number in these two linguistic aspects, recreating in his literary translation "more learned, more intellectually sophisticated characters"[12] than Clark.

Baker's model is so influential that it has ushered in a wave of corpus-based analyses of literary translators' style that continues today.[13] In addition to investigating how the overall stylistic effects of translations have been shaped by literary translators' linguistic preferences as Baker has done, later scholars also study the effects of the translators' consistent diction (e.g., contracted forms, reporting verbs, demonstrative pronouns, culture-loaded terms, archaic words) on forming a particular style in literary translations.

Focusing on the use of contracted forms such as "can't, don't, it's, they've," and so on, between Translational English Corpus (TEC) and the imaginative texts subcorpus of British National Corpus (BNC), Maeve Olohan has found significant differences between the English used by literary translators (Bush and Dorothy S. Blair, in her case) and the English used by contemporary literary writers.[14] While the occurrence of long forms, which are often used to indicate a formal tone, is much more frequent than contractions in TEC, the use of contractions instead of long forms is an obvious trend in the BNC subcorpus.[15] Additionally, the frequent use of long forms by the translator also reveals their "subconscious processes of explicitation,"[16] making literary messages less prone to ambiguous interpretations by the reader. If the translated English literary texts are examined without referring to their original texts, Olohan claims that the "genre and narrative structure"[17] are among the two main reasons that cause these stylistic variations.

Using a bidirectional Portuguese–English parallel corpus (COMPARA) as a reference corpus, Gabriela Saldanha examines how the use of foreign words by Margaret Jull Costa and Bush in their literary translations from Spanish and Portuguese into English, respectively, has contributed to their personal literary style. The significant difference she has found between the two translators in terms of their use of source culture items shows that these two translators have different target readers in mind for their literary translations. While Costa avoids exotic foreign lexical terms that might "disturb the reading experience,"[18] Bush retains more source culture items in his literary translations to produce a less familiar and coherent reading experience for the target reader. Saldanha then goes on to argue that the difference in the choice of the translation strategies of these two translators is resulted from their "different conceptualizations of their readership and of their role as intercultural mediators."[19]

In her methodology, Baker bases her study of the translator's style on translated texts only, without referring to their corresponding source texts. This approach has caused doubts among some scholars, who argue that the stylistic patterns observed in literary translations are resulted more "from the influences of the source texts"[20] than the translators' deliberate and conscious linguistic choices. Thus they contend that parallel corpus, which consists of digital collections of source texts and their originals, should be used instead of comparable corpus in reconstructing the stylistic distinctiveness of literary translators' work. Typically, in this parallel model, one literary text is investigated with its multiple translations by different translators to highlight the different stylistic choices by translators.

In the meantime, Charlotte Bosseaux investigates the translations' point of view, including deixis, modality, transitivity, and free indirect discourse, in two French translations of Virginia Woolf's novel *The Waves* from a parallel corpus perspective. Her corpus statistics are derived from the textual comparison, revealing the translators' departure from the distinctive perspective from which the six characters deliver their soliloquies. Her findings indicate clearly "the translator's voice or discursive presence"[21] in the translation process. This observation is also echoed in a series of studies done by Marion Winters on the difference of translators' style as shown in two German translations of F. S. Fitzgerald's novel *The Beauty and Damned* by Hans-Christian Oeser and Renate Orth-Guttmann.

The keywords Winters focuses on in her corpus-based study in relation to translators' style include model particles, foreign words, code-switches, and speech-act report verbs, all of which are found to be different in terms of the frequency of usage between these two translators. In her quantitative analysis of the translation of modal particles, she finds that these microlevel linguistic choices has led to a stylistic change on the macrolevel of the novel to the extent that "Orth-Guttmann provides a character study while Oeser focuses on societal issues."[22] In a recent study,[23] by using the corpus processing software Wordsmith Tools and TetraplaX, Winters studies the literary style of the German writer Natascha Wodin as reflected in her two autobiographical novels *Einmal lebt ich* and *Die glaserne Stadt* and their English translations. Her corpus analysis of the sentence length, repetition in both source texts and target texts identifies significant difference of the use of these syntactic features in the originals and the translations, despite the fact that all of these texts are actually produced by the same writer/translator. Compared with the extensive use of long sentences and repetition to emphasize the intensive feeling of the autobiographical narrator "I" in the original, Wodin as a translator uses much shorter sentences and creates more pauses by using additional

sentence markers, such as full stops, and exclamation and question marks in her translations. The overall stylistic effect of these syntactic markers, which indicate "the interruption of the stream of consciousness"[24] of the narrator, is a less intense personal narrative than the originals.

With the parallel corpus model for the study of translator's style gaining more ground around the world, some scholars in China also use this corpus methodology to carry out case studies on translator's style in English translations of Chinese literary works. By computing and comparing the STTR and the average sentence length in the two translations of the classical Chinese novel *Honglou Meng* (The Dream of the Red Chamber), Li et al.[25] investigate the stylistic difference between their translators: David Hawkes and John Minford versus Xianyi Yang and Gladys Yang. It is found that although more words are used in Hawkes and Minford's version, the STTR, which indicates the range and variety of vocabulary used, of their version is lower than that of Yang's. These stylistic differences, according to the authors, are the direct results of their different translation strategies, which were "affected by the social, political and ideological milieu in which they lived and worked."[26] In a similar vein, Liu et al.[27] analyze the translators' style in the four English versions of the same novel. However, their corpus statistics regarding the type-token ratio, lexical density, average word length, and mean sentence length reveal that there is no significant difference among them, indicating homogeneity in terms of macrolevel stylistic effects across these four translations.

Recent decades have also witnessed an increasing trend of combining both parallel corpus and comparable corpus in investigating the style in literary translations, with the aim to shed light on the stylistic profile of the translator from a multiple number of corpus indicators both from comparative and comparable perspectives. The style in the translations of modern Chinese literature has become a topic of particular interest in this new research paradigm. For example, adopting this integrated corpus approach, Qingwei Song[28] studies the style in Howard Goldblatt's English translations of the fiction written by Mo Yan, a renowned Chinese novelist who was awarded the Nobel Prize in Literature in 2012. Similar to other scholars, Song retrieves corpus statistics regarding STTR, mean sentence length, lexical density, part-of-speech (POS), and collocations in Goldblatt's translations and compares them with the fictional subcorpus of BNC. The statistics generally support the translation universals hypotheses, including simplification, normalization, and levelling-out, in the translations, which indicates a "de-complexification of language"[29] in the source texts to cater to the reading habits of target readership.

Indeed, the integrated model, which investigates the translator's style from the corpus results that are derived from both comparable and parallel corpus, can better reveal the causes of particular linguistic patterns of literary translation than Baker's monolingual comparable model as it brings in source texts into the picture. However, it is still a question that to what extent the comparative findings from the integrated model can be testified and repeated in other similar parallel corpora, such as those of translations of the same source texts, but in other languages. To introduce triangulation to the corpus findings from the integrated model, Juliane House proposes the concept of validation corpus, which consists of "translations from the same (. . .) genres into the opposite direction"[30] to validate the comparative results from the integrated model. With more than two types of corpora used in her research, this so-called "multiple-complex model"[31] enhances the reliability and validity of the research findings by making comparisons between multiple corpora possible.

Libo Huang exemplifies the application of this multiple-complex model by investigating the translator's style in the English translations of contemporary Chinese novels written by such prominent writers as Mo Yan, Lao She, Jia Pingwa, and Eileen Chang.[32] In his study on the style in three English translations of Lao She's *Luotuo Xiangzi* (Rickshaw Boy) by Shi Xiaojing, Jean M. James, and Howard Goldblatt, respectively, Huang first bases his analysis on the usual parameters such as STTR and mean sentence length, which are the frequent aspects under study in Baker's model, but finds that the corpus statistics are not significant enough to tell apart these translators' styles. He conducts the same analysis in a comparable reference corpus (i.e., his validation corpus), which consists of all these three translators' English translations of Chinese literary works except *Luotuo Xiangzi*, but comes up with similar results. As a last resort, Huang focuses on their translation of blended forms of speech representation on a parallel corpus of the source text and three translations and finds that these translators differ greatly in presenting characters' speeches and thoughts. Among these three translators, while Shi prefers to use "indirect speech with the third person" in the past tense to indicate a more distancing addresser-and-addressee relationship, James and Goldblatt tend to use "free direct speech with the first or the second personal pronoun"[33] in the present tense to accentuate a more intimate and closer atmosphere. His application of the multiple-complex corpus model to investigate the translation of local colorism in Jia Pingwa's works and readability in Eileen Chang's novel also yields interesting results, which proves the robustness of this corpus model.

Apparently, a full discussion on all related studies on translators' styles in literary translations is beyond the scope of this article. For the moment, suffice it to say that the earlier review of some of the representative works has already outlined the evolving trajectory of this research paradigm, through attempts to elaborate its theoretical frameworks and to include a wider variety of issues pertaining to translator's style in literary translation.

To begin with, it can be clearly seen that ever since its inception, the corpus-based model to literary translation research, which largely draws on its theoretical framework from CTS, is continuously developing "into a coherent, composite and rich paradigm."[34] Growing out of a quest for proving translation universals in literary translations, this digital approach to the study of translated literature has become familiar to translation scholars with Baker's comparable model for researching translator's style. At a later stage, an interlingual parallel model was more frequently adopted so that the influence of the source text on the translator's microlevel linguistic choices that lead to the individual styles of literary translation can be investigated. Then came an integrated model that computed corpus statistics from the target-text-oriented comparable model as well as the explanatory power offed by the source-oriented parallel model. Recently, there has been a tendency of applying the multiple-complex model to replace previous ones as the corpus-based research into literary translator's style has gradually shifted its focus from purely looking for statistically significant features in the corpus to validating and triangulating the comparative results in other comparable corpora. The increasing complexity and composite of the model for corpus-based stylistic research will no doubt lead to a more in-depth analysis of the translator's stylistic profile in literary translation.

Secondly, along with the development of a richer and more coherent research model, the field has also witnessed an expansion of its research territory in literary translation. Previously, what interested the scholars most were those linguistic parameters, including type-token ratio, mean sentence length, lexical variety, and lexical density, whose statistics could be generated by computer software directly. Nowadays, in addition to the investigation of these formal discourse markers, comparisons of the narrative features, such as point of view, speech and thought representation, reporting verbs, and modality among different corpus have also been introduced in the analysis so as to develop a more comprehensive view of the translator's distinctive style in literary translation. The wider variety of research topics has also led to a broadening of the scope of the literary works that are under study: although the corpus used in previous research is mostly composed of English translations of other non-English speaking writers' literary works in Europe, the

corpus that forms the basis of current research now is much more diversified, consisting of literary translations in different languages of authors around the world, including English, French, Spanish, Russian, Chinese, Korean, Japanese, and so on. The availability and diversity of these literary translation corpora make multiple comparisons across different languages possible, thus making the comparative results and findings more valid.

Lastly, there seems to be a consensus now that other than the exploration of translator's stylistic changes, including the shift of semantic, pragmatic, and narrative characteristics in literary translations as revealed from corpus statistics, equally important is the careful examination of translators' motivations behind these translation decisions. This is done by particularly referring translation strategies to the social, cultural, and literary context in which the translation activity takes place. In general, the explanations for the literary regularities or patterns found in translational behavior can be sought from a multiplicity of disciplines, including history, sociology, cultural studies, cognitive linguistics, and so on. This move to "the explorative–interpretative paradigm"[35] in exploring the style in literary translation can be seen as an attempt to combine both qualitative and quantitative inquiries and to highlight the interdisciplinary nature of this line of research.

DH Approaches to Literary Translation: The Current State of the Art

Along with the rapid technological development, especially in the field of information technology and computer science in the last decade, DH has also witnessed the application of a number of new digital methods in the scholarship and further expanded its research territory by engaging these digital tools in the study of different subjects of humanities, thus fostering new interdisciplinary areas such as digital archaeology, digital history, digital literary studies, digital media studies, digital musicology and digital sociology, and so on.

In the study of translated literature, although it is still too early to declare the birth of the discipline of digital literary translation, the field has already seen fresh attempts at building purposefully designed corpora related to special types of literary translations (e.g., poetry translation, pseudotranslation), or translations of a particular period (e.g., Chinese literary translations during the late Qing period).

For the corpora used for investigating literary translator's style mentioned in the previous section, most of them are digital collections of

contemporary literary texts as well as their translations, which are generally easier to obtain online compared with the literary works created and translated several decades ago. Moreover, the use of the modern works can quicken the text processing procedures for building a corpus, as there are usually fewer problems occurring to the recognition and proofreading processes of these texts for later retrieval by means of corpus tools. By now, attention has also been paid to the construction of databases that dates back to the beginning of the twentieth century or maybe even earlier. For instance, Nason Anran Cao explores the resources and tools for building a corpus of translated literary texts during the late Qing and Republican period (around 1870–1930), meticulously evaluating the success rate of four most frequently used Optical Character Recognition (OCR) software for recognizing vertically aligned traditional Chinese characters, which were the printing norms for all literary works back then. He concludes that a tailor-made OCR software with extra functions such as "adding more fonts to the character set or dealing with punctuations on the side"[36] should be developed for building a large-scale digitization project on this topic in the future.

Same as Cao, James St. André has also evaluated the usefulness of the existing tools for developing a purposefully built corpus. This time he wants to find out the usefulness of the large, preexisting online corpora for building translation corpora and comparable corpora of pseudotranslations (including those works claimed to be literary translations) produced centuries ago. His review of the six online corpora, including subscribed database Eighteenth Century Collections Online (ECCO), partially subscribed databases Early English Books Online (EEBO) and HathiTrust, as well as three open-access databases, Project Gutenberg, Google Books, and Internet Archive,[37] shows that despite their easy accessibility online, none of them can be used for searching statistical parameters such as type-token ratio, average sentence length, or performing POS tagging by researchers. This means that most researchers need to design and compile their own corpus in accordance with their research agendas and goals, such as developing a specialized corpus for researching special type of literary translation.

Another development in applying DH to the study of literary translation at the present time is the application of the new digital techniques such as data mining, data visualization, geographic information system (GIS), text mining, and so on, to explore and extract statistically significant information (either in the form of figures or visual elements) about the patterns and trends from a large number of written sources that might help us to discover new knowledge about the activity of literary translation of a particular translator or during a designated period of time.

For instance, by using a hybrid digital methodology, including the use of Tableau—an interactive data visualization software, David Kim[38] tracks all the translations of Goethe's first world literary text *Die Leiden des jungen Werthers* (The Sorrows of Young Werther) in the world and delineates "the many routes by which the novel travels" from Germany to other nations. Unsatisfied with Moretti's "distant reading" approach that only favors the study of "units that are much smaller or much larger than the text,"[39] Kim instead performs close reading of some key texts related to Goethe's famous pronouncements of *Weltliteratur* (world literature) and traces the possible source of the unnamed Chinese novel that inspired his idea in the first place. Based on his back-and-forth interpretations of the results of network visualization and close reading, Kim proposes that Jean-Pierre Abel-Rémusat's French translation *ou Les Deux Cousines*, which was adapted from the Chinese Novel *Iu-kiao-liI* (玉娇梨), was the most possible influence for Goethe's concept of world literature.[40] His finding shows that in order to differentiate the intricate relationships between all the related works, authors, and literary thoughts related to this concept, one needs to combine both qualitative and quantitative analyses and methods in the study of literary translation.

Data visualization, which refers to the use of graphics (such as line charts, bar chats, pie charts, points) to visually represent quantitative findings out of databases, is one of the recent platforms used by digital humanists to "analyze complex social, cultural, and historical dynamics."[41] In her use of the data visualization tool Gephi, Michelle Jia Ye[42] maps out all the translations (including those short texts appearing as in-text notes and citations), their sources, and the adjacent published texts in *New Youth*—a prominent Chinese magazine that published translations of foreign literary classics at the beginning of the twentieth century—to find out whether the selection criteria for the works to be translated are consistent for the first four volumes. Her study of the text map, which uncovers "signs of randomness, discontinuity and rupture"[43] in terms of the types and homogeneity of the materials to be translated, challenges the existing belief that the translations in the magazine are strategically and purposefully chosen to initiate a kind of new literature in China.

Similarly, David Andolfatto[44] uses word-cloud generator—a data visualization tool—to graphically examine word frequencies in the different English translations of a classical Chinese poem "Luohua" 落花 (Falling Flowers) by the renowned Chinese poet Li Shangyin in the Tang dynasty. The word clouds he has generated from these multiple English translations highlight a wide space of interpretations, ranging from convergent to divergent, by translators of different nations and periods. His findings

not only "lead to a more nuanced understanding"[45] of this eight-line poem, but also highlight the subjectivity of the translator, who has attempted the translation.

In general, the corpora used in parallel model for investigating translator's style are either aligned at paragraph level or sentence level. But in Stacey Triplette et al.'s[46] study of Robert Southey's 1806 translation *Amadis of Gaul* and its original text, they succeed in aligning the source text at clause level by applying the method of TEI XML encoding. This new DH method enables Triplette and her coresearchers to examine closely Southey's different treatments of the language of emotion in his translation, which range from retention, compression, omission, and adjustment. One of their interesting findings is that Southey actually tends to tone down or attenuate the original descriptions of emotion despite the fact that he himself is a Romantic poet well known for his compassionate writings.

The studies reviewed earlier amply demonstrate how the new digital methodology introduced by DH can help to unravel the intricate social, cultural, and literary networks in which the activities of literary translation are conducted. The corpus statistics, text mining results, and digital visualizations offered by DH approach provide us with a more systematic and graphic understanding of the various interlocking factors at work during the production process of literary translation than traditional methods (such as close reading), thus opening up new ways of interpreting and analyzing translated literature. It is foreseeable that DH methods will continue to inspire research on literary translation with systematic application and analysis of purposefully collated digital resources either on the background, the source texts and target texts, or the related works done by writers, translators, or scholars across different periods of time, an almost insurmountable task that would not have been accomplished without the various computing methods introduced by DH.

Conclusion

Although the application of DH approaches unmistakably yields new insights and findings related to the study and interpretation of literary translation, which results in a complex process involving the recreation of *literariness*[47] in the target text, some scholars raise the question whether the computer-assisted visualization and reading practices can really provide a full picture of the networking in which this recreation is performed, since there are always potential risks of "the incompleteness of searchable data or

the retrieval of irrelevant data"[48] in a given research project as well as simple hypotheses inferred from computing statistics that can also be confirmed intuitively. In our opinion, the quantitative trend of using big data to find patterns and relations brought about by DH should be welcomed, given that it can encourage scholars to base their evaluations of literary translations on consistent textual facts, rather than on their conjectural speculations only. Additionally, it is exactly the lack of explanatory power of the digital data that calls for a more active role to be played by literary translation researchers. On the other hand, however, instead of entirely relying on digital surrogates, it is advisable for translation scholars to integrate corpus statistics, computing data and visualizations into rigorous interpretations of literary translations as facts of target culture literary system, and to differentiate the complex relationships involved in the system by means of close reading and contextualization. Only by doing so can we benefit the most from the novel research and computational tools offered by DH and gain a better understanding of literary translation as a creative rewriting exercise in today's data-driven society.

YIFENG SUN is chair professor of Translation Studies in the Department of English and senior research fellow in the Institute of Advanced Studies in Humanities and Social Sciences at the University of Macau, as well as honorary professor at the University of Queensland. He is editor-in-chief of *Babel: International Journal of Translation*, and former vice president of the International Association for Translation and Intercultural Studies. He is the author of *Translational Spaces* (in press), *Translating Foreign Otherness* (2018), *Cultural Translation* (2016), *Perspective, Interpretation and Culture* (2nd edition, 2006), *Cultural Exile and Homeward Journey* (2005), and *Fragmentation and Dramatic Moments* (2002), and editor or coeditor of *Translating Chinese Art and Modern Literature* (2019), *Translation and Academic Journals* (2015), and *Translation, Globalisation and Localisation* (2008).

DECHAO LI is associate professor of Department of Chinese and Bilingual Studies, The Hong Kong Polytechnic University. He also serves as the chief editor of *Translation Quarterly*, a journal published by the Hong Kong Translation Society. His main research areas include corpuses-based translation studies, empirical approaches to translation process research, history of translation in the late Qing and early Republican periods, and PBL and translator/interpreter training. He has published over sixty articles in journals such as *Perspectives: Studies in Translation Theory and Practice*, *The Translator*

and Interpreter Trainer, *Chinese Translators Journal*, *Modern Foreign Languages*, *Foreign Language Teaching and Research*, and *Journal of Foreign Languages* as well as some book chapters published by Routledge, Springer, and Wayne State University Press.

Notes

1. Marija Dalbello, "A Genealogy of Digital Humanities," *Journal of Documentation* 67, no. 3 (2011): 481.

2. Ning Wang, "Introduction: The Interactive Relations between Science and Technology and Literary Studies," *Comparative Literature Studies* (same issue).

3. Qinglong Peng, "Digital Humanities Approach to Comparative Literature: Opportunities and Challenges," *Comparative Literature Studies* (same issue).

4. Jessica DeSpain and Jennifer Travis, "Introduction: Digital Humanities and the Nineteenth-Century American Literature Classroom," in *Teaching with Digital Humanities: Tools and Methods for Nineteenth-Century American Literature*, ed. Jessica DeSpain and Jennifer Travis (Urbana, Chicago; Springfield: University of Illinois Press, 2018), ix–xxviii.

5. Susan Hockey, "The History of Humanities Computing," in *A Companion to Digital Humanities*, ed. Susan Schreibman, Ray Siemens, and John Unsworth (Oxford: Wiley Blackwell, 2004), 4–13.

6. Anne Burdick, Johanna Drucker, Peter Lunenfeld, Todd Presner, and Jeffrey Schnapp, *Digital Humanities* (Cambridge: MIT Press, 2012), 123.

7. Gideon Toury, *Descriptive Translation Studies and Beyond* (Amsterdam; Philadelphia: John Benjamins Publishing Company, 1995), xii.

8. Ibid., 23.

9. Mona Baker and Gabriela Saldanha, ed. *Routledge Encyclopedia of Translation Studies* (London; New York: Routledge, 2020), 379.

10. Cf. Mona Baker, "Corpus Linguistics and Translation Studies: Implications and Applications," in *Text and Technology: In Honour of John Sinclair*, ed. Mona Baker, Gill Francis, and Elena Tognini-Bonelli (Amsterdam; Philadelphia: John Benjamins Publishing Company, 1993), 233–50.

11. Cf. Mona Baker, "Towards a Methodology for Investigating the Style of a Literary Translator," *Target* 12, no. 2 (2000): 241–66.

12. Ibid.

13. Cf. Libo Huang, *Style in Translation: A Corpus-Based Perspective* (Heidelberg: Springer; Dordrecht; New York; London: Shanghai Jiao Tong University Press and Springer, 2015), 26.

14. Cf. Maeve Olohan, "How Frequent Are the Contractions? A Study of Contracted Forms in the Translational English Corpus," *Target* 15, no. 1 (2003): 59–89, 59.

15. Ibid., 84.

16. Ibid., 85.

17. Ibid., 84.

18. Gabriela Saldanha, "Style of Translation: The Use of Foreign Words in Translations by Margaret Jull Costa and Peter Bush," in *Corpus-Based Translation Studies: Research and Applications*, ed. Alet Kruger, Kim Wallmarch, and Jeremy Munday (London; New York: Continuum, 2011), 237–58, 256.

19. Ibid., 257.

20. Huang, *Style in Translation*, 26.

21. Charlotte Bosseaux, "Translating Point of View: A Corpus-Based Study," *Language Matters* 35, no. 1 (2004): 259–74, 259.

22. Marion Winters, "Modal Particles Explained: How Modal Particles Creep into Translations and Reveal Translators' Styles," *Target* 21, no. 1 (2009): 74–97, 94.

23. Marion Winters, "The Case of Natascha Wodin's Autobiographical Novels: A Corpus-Stylistics Approach," in *The Palgrave Handbook of Literary Translation*, ed. Jean Boase-Beier, Lina Fisher, and Hiroko Furukawa (Cham: Palgrave Macmillan, 2018), 145–66.

24. Ibid., 161.

25. Defeng Li, Chunling Zhang, and Kanglong Liu, "Translation Style and Ideology: A Corpus-Assisted Analysis of Two English Translations of *Hongloumeng*," *Literary and Linguistic Computing* 26, no. 2 (2011): 153–66.

26. Ibid., 160.

27. Zequan Liu, Chaopeng Liu, and Hong Zhu, "Honglou Meng sige yingyiben yizhe fengge chutan" [An Exploration of the Translators' Styles of Four English Versions of *Honglou Meng*], *Chinese Translators Journal* 32, no. 1 (2011): 60–64.

28. Qingwei Song, *Mo Yan xiaoshuo yingyi fengge yanjiu* [Style in the English Translations of Mo Yan's Fiction: A Corpus-Based Study] (Jinan: Shandong University Press, 2014).

29. Federico Zanettin, *Translation-Driven Corpora* (London; New York: Routledge, 2012), 13.

30. Juliane House, "Using Translation and Parallel Text Corpora to Investigate the Influence of Global English on Textual Norms in Other Languages," in *Corpus-Based Translation Studies: Research and Applications*, ed. Alet Kruger, Kim Wallmarch, and Jeremy Munday (London; New York: Continuum, 2011), 187–208, 191.

31. Huang, *Style in Translation*, 11.

32. Cf. Ibid.

33. Ibid., 55.

34. Sara Laviosa, "The Corpus-Based Approach: A New Paradigm in Translation Studies," *Meta* 43, no. 4 (1998): 474–79, 474.

35. Cf. House, "Using Translation and Parallel Text Corpora," 206.

36. Nason Anran Cao, "Resources and Tools for Corpus Compilation of Translated Literary Texts in Late Qing and Republican Period," *Journal of Translation Studies* 2, no. 1 (2018): 153–68, 166.

37. James St. André, "A Cost-Benefit Analysis of Using Online Corpora for Translation Research," *Journal of Translation Studies* 2, no. 1 (2018): 127–52, 128.

38. David D. Kim, "Networking the Missing Linkage in Goethe's *Weltliteratur:* Orientalism and Jean-Pierre Abel-Rémusat's Sinology," *Journal of Translation Studies* 2, no. 1 (2018): 5–34.

39. Franco Moretti, "Conjectures on World Literature," *New Left Review* 1 (2000): 54–68, 57.

40. Cf. Kim, "Networking the Missing Linkage in Goethe's *Weltliteratur*," 16.

41. Burdick et al., *Digital Humanities*, 18.

42. Cf. Michelle Jia Ye, "Expanding Translation: A Text Map of New Youth (1915–1918)," *Journal of Translation Studies* 2, no. 1 (2018): 59–106.

43. Ibid., 79.

44. Cf. Lorenzo Andolfatto, "Thick Translation through Word-Clouds; or, an Educated Form of Tasseography," *Journal of Translation Studies* 2, no. 1 (2018): 107–25.

45. Ibid., 113.

46. Cf. Stacey Triplette, Elisa Beshero-Bondar, and Helena Bermúdez Sabel, "A Digital Humanities Approach to Cultural Translation in Robert Southey's *Amadis of Gaul*," *Journal of Translation Studies* 2, no. 1 (2018): 35–58.

47. Cf. Cecilia Rossi, "Literary Translation and Disciplinary Boundaries: Creative Writing and Interdisciplinarity," in *The Routledge Handbook of Literary Translation*, ed. Kelly Washbourne and Ben Van Wyke (London; New York: Routledge, 2019), 42–57, 43.

48. Kim, "Networking the Missing Linkage in Goethe's *Weltliteratur*," 14.