

Original Paper

Translation Studies from the Perspective of Corpus Translation Studies and Digital Humanities

Jianghui Xu¹

¹ Nanjing University of Information Science & Technology, Nanjing, Jiangsu Province, 210044, China

Received: June 26, 2023

Accepted: July 24, 2023

Online Published: July 27, 2023

doi:10.22158/jetss.v5n3p78

URL: <http://dx.doi.org/10.22158/jetss.v5n3p78>

Abstract

As a translation research method and paradigm, corpus translation studies have gradually gained attention from the translation community since their introduction to China in the 1970s. With the rapid development of information technology and artificial intelligence, corpus translation studies have continuously made new progress. With the support of digital humanities technologies such as big data, artificial intelligence and cloud computing, the development of corpus translation presents new features and trends. This article explores the application of corpus translation studies in translation studies from both theoretical and practical perspectives. At the theoretical level, this article believes that corpus translation studies have significant implications for the paradigm shift, theoretical innovation, and disciplinary development of translation studies. At the practical level, this article points out that corpus translation studies have strong data support, which can improve the efficiency and accuracy of translation research. Based on this, this article proposes that future translation research should focus on data-driven research paradigms, attach importance to the important role of data in translation, and build a theoretical model of translation studies on this basis, in order to provide theoretical guidance for comprehensively improving the quality of translation research.

Keywords

Corpus Translation Studies, Digital humanities, Data driven, Translation Theory Model

1. Introduction

Corpus translation studies and digital humanities have significant and extensive application value in the field of translation research. They provide new theoretical perspectives, methods, and tools for translation research, and have played a positive role in promoting the translation process, translation quality assessment, and the development of translation theory. Firstly, corpus translation studies are a research method based on large-scale language data, which uncovers translation phenomena and laws

in the translation process by collecting, organizing, and analyzing corpora from the real world. Corpus translation studies use corpus as a means of studying translation phenomena or activities (Hu, 2012, p. 60). Corpus translation studies provide a wealth of authentic language usage information, enabling researchers to gain a deeper understanding of the differences between different languages, terminology usage in professional fields, common translation problems, and more. This data-driven approach can help researchers improve translation quality, optimize translation processes, and provide guidelines and strategies to guide translation practice. Secondly, digital humanities includes the research and analysis methods of Humanities using computer and digital technology. In the field of translation research, digital humanities provides tools and techniques for processing and analyzing large amounts of text data, which can help researchers discover and mine information and patterns hidden behind language. For example, using Text mining technology, researchers can analyze translation choices, terminology use, style characteristics, etc. in large-scale corpora, thus revealing differences between different cultures and languages, and providing more accurate and effective translation strategies for cross-cultural communication.

The application of corpus translation studies and digital humanities in the field of translation research has broad value. First, they can support the development and improvement of automated translation tools such as Machine translation and computer-aided translation. By analyzing and comparing bilingual data in large-scale corpora, the performance of translation systems can be optimized, and translation quality and efficiency can be improved. Secondly, corpus translation studies and digital humanities methods can also be used for translation education and training, helping students and professional translators improve their translation skills and decision-making abilities. In addition, they can also be applied in fields such as literary translation, legal translation, and business translation, providing more accurate, consistent, and fluent translation results.

In summary, the importance and application value of corpus translation studies and digital humanities in the field of translation research are increasingly prominent. By utilizing large-scale corpora and digital technology, researchers can deeply explore translation phenomena, improve translation quality, promote cross-cultural communication and understanding, and provide beneficial support for translation practice and education.

2. The Basic Principles and Methods of Corpus Translation Studies

Corpus linguistics is a new discipline developed in the past 20 years. On the one hand, it provides new research means, on the other hand, it reflects the theoretical thinking on language based on a large number of language facts in corpus. The same applies to corpus translation studies, which are even younger, with only over a decade of research history, including applied research at the methodological or instrumental level and theoretical research on the abstraction of translation features (Wang, 2006, p. 8). Corpus translation studies are translation studies methods based on corpus data, mainly analyzing and applying large-scale bilingual or multilingual corpora to research and improve translation quality.

The basic principles and methods can be summarized as follows:

(1) Data driven, corpus translation studies emphasize the importance of data in translation. It ensures that the translation results are more in line with practical application needs by collecting and utilizing large-scale real bilingual corpus data.

(2) Parallel corpora are used in translation studies, which are corpora that contain both source and target language texts. These parallel corpora can provide sentence level correspondence and provide basic data support for cross language translation.

(3) Alignment and correspondence establishment: Alignment is one of the core steps in corpus translation studies. By using automatic or manual alignment operations, correspondence between sentences in the source and target languages is established, thereby establishing correspondence between different languages.

(4) Translation memory: Translation memory is one of the important tools of corpus translology. It improves translation efficiency and consistency by storing and reusing previously translated sentences and phrases.

(5) Term extraction and management: Corpus translation studies can extract and manage professional terms from corpora through automated methods, ensuring consistency of terms during the translation process.

(6) Statistical Analysis and Machine Learning: Corpus Translation is based on statistical analysis and machine learning techniques, which improve translation quality, optimize translation algorithms and models by analyzing and modeling large-scale corpora.

(7) Cross language information retrieval and application: Through corpus translation studies, cross language information retrieval and application can be achieved, that is, using source language queries to search for relevant texts in the target language, expanding the scope of obtaining text information.

In general, corpus translology, based on data, uses parallel corpora, Translation memory and other tools, combined with statistical analysis and machine learning technology, to improve translation quality, improve translation efficiency, and support cross language information processing and application.

3. Translation Studies from the Perspective of Digital Humanities

Digital humanities, also known as humanistic computing, is “an interdisciplinary research field formed by integrating Humanities knowledge, data collection and analysis technology, network infrastructure and algorithm models” (Hu & Hei, 2020, p. 5). The application of digital humanities in translation research involves multiple aspects. Firstly, in terms of resource construction and organization, digital humanities technology can help researchers build and organize large-scale corpora, including parallel corpora and monolingual corpora. These corpora provide a rich data foundation for translation research. In addition, in terms of automatic alignment and comparison, digital humanities tools can automatically align corresponding texts between different languages, thereby finding their interrelationships. This

alignment and comparison helps researchers analyze and compare cross language texts, providing important support for translation research. Of course, as for the improvement of Machine translation, digital humanities technology can be applied to the training and improvement of Machine translation system. By using large-scale corpus for machine learning and model optimization, the performance and quality of Machine translation systems can be improved. Finally, in terms of innovative research methods, digital humanities provide new research methods and means for translation research. For example, Text mining and visualization technology can be used to discover the rules and trends in the process of translation, and to explore the essence of translation strategies and decision-making. In summary, the application of digital humanities in translation research has provided researchers with abundant resources and tools, promoting the development of translation theory and practice.

4. A Cross Disciplinary Study of Corpus Translation Studies and Digital Humanities

Since Baker (1993) first discussed the application of corpus technology in translation research, translation research from the perspective of digital humanities has developed rapidly. Corpus translation studies and digital humanities methods are two important branches in the field of translation research, which can be organically combined through cross research to promote the in-depth development of translation research. Corpus translation studies utilize large-scale parallel corpora for data-driven research, revealing translation phenomena and patterns through statistical analysis and machine learning methods. It focuses on research on translation units, translation difficulties, translation strategies, and provides empirical data support, which has had a positive impact on translation theory and practice. Digital humanities method is to use computer and digital technology to process and analyze data and text of Humanities. It includes Text mining, Natural language processing, image recognition, machine learning and other technologies, which can help researchers acquire, organize and analyze large-scale text and corpus data. In cross research, corpus translology can draw on the technical means of digital humanistic methods, such as using Text mining and machine learning technology to mine the potential laws and features in parallel corpora, so as to better understand and interpret the translation process and results. On the other hand, digital humanities methods can enrich their application scenarios and provide more accurate and comprehensive research results through the theoretical foundation and data resources of corpus translation studies. Cross disciplinary research can also promote the application of corpus translation studies and digital humanities methods in practice. For example, using corpus translation methods to analyze translation quality and evaluate model effectiveness, and utilizing the data processing capabilities of digital humanities methods to monitor and improve the translation process. In addition, combining the research results of the two can also promote the development of translation tools and the optimization of automated translation systems. In short, the cross research between corpus translation studies and digital humanities methods provides new ideas and methods for translation research, which can support a deeper and comprehensive understanding of translation phenomena and laws, and bring more innovation and improvement to translation practice.

5. Conclusion

In summary, China has achieved remarkable results in translation research from the perspective of digital humanities, and the academic community's understanding of translated texts and translation processes is becoming increasingly objective and comprehensive (Hu, 2018, p. 25). With the support of artificial intelligence technology and big data technology, corpus translation is gradually becoming a science and developing to the application level. Under the data-driven research paradigm, translation research should pay more attention to the role of data in translation, and construct translation theoretical models based on data to improve the efficiency and accuracy of translation research. Chinese translation researchers should fully leverage the advantages of corpus translation studies and digital humanities, actively draw on advanced theories and methods from abroad, and continuously promote the development of Chinese translation research. At the same time, translation researchers should pay attention to the exchange and cooperation between domestic scholars and international peers, draw on advanced foreign experience, combine with China's national conditions and translation practice, and continuously improve and develop the corpus translation theory system with Chinese characteristics.

References

- Baker, M. (1993). Corpus linguistics and translation studies: Implications and applications. In M. Baker, G. Francis, & E. Tognini-Bonelli (Eds.), *Text and Technology in Honor of John Sinclair*. Amsterdam: John Benjamins. <https://doi.org/10.1075/z.64>
- Baker, M. (1996). Corpus-based translation studies: The challenges that lie ahead. In H. Somers (Ed.), *Terminology, LSP and Translation: Studies in Language Engineering*. In Honour of Juan C. Sager. Amsterdam: John Benjamins. <https://doi.org/10.1075/btl.18.17bak>
- Hu, K. B. (2012). Corpus Translation Studies: Connotation and Meaning. *Foreign Language*, (5), 59-70.
- Hu, K. B. (2018). Progress and Prospects of Translation Research from the Perspective of Digital Humanities. *China Translation*, 39(06), 24-26.
- Hu, K. B., & Hei, Y. (2020). Translation Studies from the Perspective of Digital Humanities: Characteristics, Fields, and Significance. *Chinese Translation*, (2), 5-15.
- Tan, S. R. (2016). Fu Liping Empirical Research on Translation Process: Achievements, Shortcomings, and Implications. *Oriental Translation*, (5), 23-27.
- Wang, K. F. (2006). Corpus Translation Studies - A New Research Paradigm. *Chinese Foreign Language*, (3), 8-9.
- Yang, H. Z. (1993). Corpus linguistics and Machine translation. *Journal of Shanghai Jiao Tong University (Philosophy and Society Edition)*, (1), 98-11.