Research on the Relationship between Corpus Linguistics and Translation

Shan Da

Xi'an Peihua University, Shaanxi 710125, China

331849629@qq.com

Abstract. The corpus is the database responsible for the storage of language materials. As a large-scale electronic text library, researchers can conduct related language theory and applied research by using computer analysis tools to. It is widely used to provide relevant statistics or case studies for many fields such as language teaching, traditional language research, natural language processing, and lexicography. Linguistics is a branch of corpus linguistics, and it is inseparable from the corpus. The research direction of this paper mainly focuses on the relationship between corpus linguistics and translation and puts forward some viewpoints on it.

Keywords: corpus; linguistics; translation.

1. Introduction

The early corpus was formed before the mid-1950s and has been developed for more than half a century. The study of corpus is no longer only serving certain special groups under the influence of continuous advancement in computer technology; even its development speed is affecting more and more fields. Corpus linguistics is a discipline that develops new research techniques and methods based on a corpus. The corpus linguistics opens a new chapter in language research. Corpus translation is also a new subject; its translation research ways are carried out through corpus, and combined with corpus linguistics. The translation based on corpus linguistics has obtained a new theoretical application platform, which changes the translational normality of traditional concept and constraints. This paper will explore the construction of corpus linguistics in the real state, and try to sort out the development and application of corpus translation.

2. The Concept an Development of Corpus Linguistics

The corpus is also called as material, which is a natural occurrence of a linguistic material and consists of spoken and written language, it can be used both as a starting point for language description and as a measure to prove the language authenticity. The corpus is a database of linguistic materials that is stored and reasonably arranged. This kind of linguistic material, either written or verbal, occurs naturally.

Corpus linguistics emerged as a new discipline at the end of the 19th century; this period was regarded as the first stage of research, its value is reflected in the service of text retrieval, sampling, analysis and arrangement. In the 1960s, corpus linguistics reached the second stage. In modern China, corpus linguistics has been highly valued by more and more scholars and experts, including parallel corpora, corpora for special purposes, special corpora for oral and written feature learning, and other corpora have been developed. The third development stage is in the mid-1990s, and the research direction has shifted from singleness to diversification, including research and teaching in foreign languages, translation theory and practice, discourse error analysis, language processing and second language acquisition and so on.

After more than half a century of development, corpus linguistics has become a typical interdisciplinary field combined with multiple disciplines. Many discipline studies that have been involved before have achieved great results. The importance of corpus in language research is recognized by more and more people. The construction of a large number of corpora has become a key factor in the rapid development of many language-based research projects; corpus linguistics continues to maintain an unparalleled role in the study of language.

How to define the status of corpus linguistics is a question. It is believed that linguistics is an independent, emerging linguistic sub-discipline, and others think that research methods based on corpus can enhance the status of linguistics and define its nature. The second is the problem of building a corpus. The construction of the corpus is related to the format and copyright of the text, in addition, the coding and marking systems of corpus are different, which also has a great hindrance to the research.

The advance of science and technology leads the rapid development of network information technology, and the storage capacity of the corpus is no longer a problem. The variety of corpora is becoming more and more abundant. The corpus is more closely integrated with many disciplines such as lexicography, artificial intelligence, and computational linguistics to achieve more effective and smoother language research.

3. Understanding of Corpus Translation Studies

Corpus translation studies is also a discipline based on corpus linguistics, it has special place with its own, first, and it is a behavior of corpus research based on its own knowledge background and consciousness at a specified time and place. The extensive translation comes from the general direction of a large number of translated texts, but not the objective laws; secondly, corpus translation studies are based on the process and law of two languages and mutual conversion, it is necessary to rely on a more complicated bilingual corpus, finally, corpus translation studies is to develop a translation corpus while leading the full translation corpus.

The corpus was discovered for translation studies in the 1980s. Modern translation studies consist of two parts: the first part refers to the comparison and collation analysis of language phenomena related to translation, it is completed by using traditional manual collection of original text and translated text, the second part is the practice of various translation products in the translation process. As the corpus becomes more prominent in the field of translation studies, more researchers have understood the close relationship between the two, which inevitably led researchers to generate new thinking problems: Are the necessary conditions for translating texts simple, clear and standardized? This requires more corpus-based validation of translation studies to draw a conclusion.

The development time of corpus translation studies is not very long, but it has received the attention of the majority of translation researchers. In the translation studies focusing on translational commonality research, the importance of corpus translation studies is becoming more and more obvious, and thus a series of research results with great influence are studied. Corpus translation studies are more reflected in the impact on traditional translation concepts and research methods.

4. Application Scope of Corpus Translation Studies

With the wide application of corpora in translation studies, many researchers use corpus analysis tools to search corpus data, data statistics and analysis are conducted on indicators that reflect language characteristics, such as type token ratio, lexical density, collocation pattern, average sentence length, sentence pattern, etc., then examine the vocabulary diversity and information load in the translated text. In addition, some researchers combine theoretical interpretation and empirical research to explore the characteristics of translated texts and so on. These are also the main means of current corpus translation research methods. Corpus translation studies are more scientific than traditional translation research methods.

There are three types of corpora that can be used for translation studies: one is the parallel corpus; the other is the multilingual corpus; the third is the comparable corpus. The parallel corpus refers to a bilingual parallel corpus, which collects the original text of a language and its translated text in another language, and it can be said to be a translation corpus in the true sense. The multilingual corpus is the corpus that collects texts two or more languages with similar genres and texts. A comparable corpus is a corpus that collects the original text of a language and texts with similar text translated from other languages.

The establishment of comparable corpora has the most practical value for translation research, because researchers can use the comparable corpus to analyze and compare original texts and foreign translations in the same language, reveal their differences from vocabulary to style, and discover translational norms and practices in a particular historical and cultural context. It can be seen from this that the translation norm research and translational universality has a broader perspective with the help of corpus. The broad application prospects of corpus translation research methods in translation research are mainly reflected in the following aspects:

First, it is expressed in translation universals research. The translation universals research is one of the most important application areas of corpus translation research. Revealing and verifying the existence of translation universals is the greatest achievement of corpus translation research so far. Mona Beck first used parallel corpus for translation universals research, revealed the universal characteristics of translation texts with simplicity, clarity and standardization, and further pointed out that translation universals are a typical feature inherent in translated text itself. However, the current research on translation universals is only the initial result of corpus translation research; the conclusion on translation universals is inevitably restricted by the objective conditions at that time, such as the size of corpus, the depth of corpus processing, the function of related retrieval software, and the maturity of inspection and analysis methods. Limited by the above objective conditions, the corpus-based translation universality research needs to be further improved both in terms of research methods and research processes. Limited by the above objective conditions, the corpus-based translation universals research needs to be further improved in research methods and research processes.

Second, any translation activity cannot be separated from specific social and cultural background, so the translation strategies and products are bound to be restricted by language norms and social norms. The translation academics' discussion of translation norms mainly focuses on the formulation of translation product specifications and statute standards. However, traditional translation studies are basically qualitative research, and subjectivity is strong. Based on this point, the application of corpus provides a new perspective for translation norm research. One of the advantages of corpus translation research is its large amount of text corpus, based on these corpora, various normative factors affecting translation can be described and analyzed, which is more scientific and objective. Translation norms are divided into internal norms and external norms, among them, internal norms are subdivided into compulsory norms, communicative norms and operational norms, and external norms are subdivided into selection norms and expectation norms. This corpus-based translation specification study provides a paradigm for the descriptive study of translation norms. It should be pointed out that there are still many problems in the study of translation norms that need to be further explored.

Another manifestation is the translator's stylistic research. Before the rise of corpus translation studies, most of the translation style studies use typical examples from the text as the basis, the impressions and empirical reviews were mainly subjective, and it is difficult to conduct a comprehensive and quantitative study on large-scale translated texts. Mona Baker opened a new field of corpus translation research, and provided a new paradigm for the study of translator style. With the help of the corpus, the scope of the translator's stylistic research is also broader; they can study the language style and mode of the original work, translation and translator's language style and mode, the style of multiple translations of an original work, study different translators' similar treatment and different grasps of the same original text, and explore the translator's cognitive process and social and cultural orientation. In short, the translator's style is studied based on the corpus translation research method, investigate the many factors affecting the translator's style through further quantitative research, avoid subjective judgment, it is a new paradigm of the translator's stylistic research.

As a research means and method, the scope of corpus translation research is not limited to the study of translation universals, translation norms and translator style, but also involves many other fields, such as translation text vocabulary innovation, translation teaching, and interpretation research. Chen Wei explains the application potential of Translational English Corpus(TEC), points out that "the research areas that can be developed at this stage include: translation studies in translated language culture, research on ellipsis question in translation, translator's personality and gender

impact on translation, synchronic and diachronic research on the translation of literary works in different periods and in the same period,, female translation studies, how the regulation of subject culture restricts and influences the study of translation policies and strategies, study of the differences in translation techniques and strategies among different styles of writing." TEC translation corpus combines with other analogical corpuses to study the style of the text, the translation characteristics of new words, language habits and other questions. It can be seen that the application of corpus translation research is more and more extensive, and its emerging important discoveries and research results are also remarkable.

5. Advantages of Corpus Translation Studies

The corpus-based translation research method has the following three advantages: first, it can conduct fast, accurate and complex analysis by computer; second, the corpus is large in scale, including comprehensive corpus, large amount of text, and wide range of linguistic information, third, there are both quantitative analysis and qualitative functional interpretation, and the description of language is comprehensive. Moreover, corpus translation research combined with quantitative and qualitative methods, which is not restricted by the development of linguistics, but extends the research scope to social and cultural factors; it is not just a description of data, it also explains the social causes behind translation phenomena on the basis of descriptions. The advantages of corpus translation research method are unparalleled for translation training and translation teaching as well as translator's translation practice. The corpus as a teaching resource, the greatest advantage of introducing translation classroom teaching lies in the authenticity and effectiveness of the corpus. On the other hand, since the material of the modern corpus is composed of electronic versions, it can be reused by teachers and students, which make teaching independent and open, and promote students' data-driven learning. As far as the translator's translation practice is concerned, the corpus provides rich illustrations and large amount of information, which can be retrieved quickly and conveniently, the translators have a lot of leeway and strong autonomy, in addition, the material in the corpus can be extracted and edited, which greatly improve the translator's translation efficiency and quality.

6. Conclusion

Corpus-based translation studies extend the scope of translation research and introduce a new research paradigm for translation studies. Corpus translation studies have a profound impact on the entire translation research, which can help researchers to understand translation and explain translation with a broader perspective. As a research method, the corpus will continue to improve with the pace of social development, and will also provide more efficient and high-quality language material services for translation studies, thus promoting the further study of corpus translation studies.

References

[1]. Ding Shude. A Study of Western Translational English Corpus[J]. Journal of Foreign Languages, 2001, (5): pp.61-66.

[2]. Hu Xianyao. Contemporary Corpus-Based Translation Studies and Universals of Translation[J]. Shanghai Journal of Translators for Science and Technology, 2004, (4): pp.47-49.

[3]. Liao Qiyi. Corpus and Translation Studies[J]. Foreign Language Teaching and Research, 2000, (5): pp.380-384.

[4]. Wang Haiping, Hu Weihua English-Chinese Parallel Corpus and Translation Studying[J], Education Science & Culture Magazine, 2011, (10): pp.129-130.

[5]. Zhao Xiuli. Application of Corpus in Translation Studies and Translation Teaching[J], Journal of Nanchang College of Education, 2011, (7): p.179.