Original Paper

On the Translation Strategies of Human-computer Interaction Based on Machine Translation

QIN Fangfang[1] & XIANG Xiaomei[2*]

[1] School of Foreign Languages of China Three Gorges University, Yichang City, Hubei Province, China

[2] College of Humanities and Social Sciences of Xinjiang Institute of Technology, Aksu, Xinjiang, China

[*] XIANG Xiaomei, College of Humanities and Social Sciences of Xinjiang Institute of Technology, Aksu, Xinjiang, China

Abstract

With the rapid development of language service industry and artificial intelligence technology, machine translation plays a more prominent role in the translation industry. Human-computer interaction translation greatly improves the speed and quality of translation, and pre-translation editing and post-translation editing are two important links and manifestations in human-computer interaction collaborative translation. On the basis of summarizing machine translation problems, this paper proposes translation strategies including replacing well-translated terms in advance, rewriting, addition, omission, and shift via pre-editing and post-editing, which greatly improves the quality of machine translation.

Keywords

human-computer interaction translation, machine translation, pre-editing, post-editing

1. Introduction

From the beginning of the 21st century, machine translation technology has developed rapidly in China. Faced with the surge in market demands for language services, human translation or pure machine translation can no longer meet these needs. In recent years, with the emergence of neural network machine translation and the "machine translation + post-translation editing" model, the quality and efficiency of translation have been greatly improved, but there is still a certain gap compared to professional human translation. So, how to further innovate on the basis of machine translation? In other words, what else can we do to improve machine translation? In view of this, based on summarizing machine translation problems, the author proposes the "pre-translation editing + machine translation + post-translation editing" model, and concludes a series of translation strategies from pre-translation

editing and post-translation editing, thus providing a certain reference for future human-computer interaction translation research and application.

2. Research Status

The study of how to use computers to automatically convert natural languages is one of the important research fields in artificial intelligence and natural language processing. As a key technology to break through the language barrier by information transmission between different countries and ethnic groups, machine translation is of great significance for promoting ethnic unity, strengthening cultural exchanges and promoting foreign trade (Liu, 2017, p. 1144).

In the 17th century, Descartes and Leibniz proposed the idea of using machine dictionaries to overcome language barriers. IBM's 701 computer automatically translated 60 Russian sentences into English, the first machine translation in history. Machine translation has experienced a long and tortuous development process, which is generally divided into the following four stages: the inception period (1949-1960), the setback period (1960-1967), the recovery period (1967-1990) and the new period (1990-present) (Gao & Zhao, 2020, pp. 97-98).

Although the quality of machine translation has increased, it is still not comparable to high-quality human translation. Pure machine translation cannot guarantee the translation quality, let alone meet the increasing demand for language services. Therefore, the function of human-computer interaction translation has become increasingly prominent with the pre-editing and post-editing as its important manifestations. "Pre-translation editing is mainly aimed at the source text. It refers to the targeted modification and editing of the texts or documents that need to be translated before machine translation to improve the quality of machine translation output, and post-translation editing means to modify and edit the original output of machine translation. Its purpose is also to improve the quality of machine translation through human-machine interaction" (Feng & Gao, 2017, p. 63).

At present, in-depth research on machine translation has been carried out in China. Post-translation editing is also a new hot research topic. The post-translation editing is an important manifestation of human-computer interaction, a new business growth point for language service companies, and the development direction of translation services in the future (Cui, 2014, p. 70). At the same time, the research on pre-translation editing has not been fully carried out. Some researchers believe that the advantages of pre-translation editing are not obvious, and there are defects such as large workload and low efficiency. However, the author believes that proper pre-editing can increase the quality and acceptability of machine translation. Moreover, the enthusiasm of scholars for pre-translation editing has gradually increased in recent years, and a series of researches have been carried out around it. It can be clearly found out from CNKI data that the number of related journals published about pre-translation editing since 2018 has increased year by year. Therefore, the author proposes the "pre-translation editing + machine translation + post-translation editing" model, hoping to improve the machine's understanding of the source text to reduce the workload of post-translation editing, and promote translation efficiency

44

through pre-translation editing. However, even after pre-editing and machine translation, there are still some problems in the translation, and in this case, post-editing is necessary.

3. Common Problems of Machine Translation

Although machine translation has greatly improved translation efficiency, there are also some unavoidable problems in its original output, which make the translation quality unsatisfactory, such as follows:

3.1 Terms Inconsistency

A term is used to express the exact concept of a particular thing. The "inconsistency of terms" means in the process of translation, namely, conversion from the source language to the target language, one term of the source language has different expressions, but the multiple expressions in the source text for the same thing are translated into different versions in the target text by machine. The difficulty for computer to analyze Chinese language lies in the fact that the same part of speech in Chinese serves as multiple grammatical components without morphological changes, and the formation principles of Chinese sentences are basically the same as those of phrases (Guo & Wang, 2017, p. 78). Especially for large texts, it is prone to cause terms inconsistency by machine translation because of the context or the different collocations of the same term.

3.2 Improper Segmentation of Punctuation Marks

Punctuation marks are an integral part of a written language. At present, the punctuation marks used in Chinese are formulated based on the English punctuation system. They do not only retain most of the main characteristics of English punctuation, but also shows the characteristics of Chinese language. Therefore, there are some differences between Chinese and English punctuation marks, which are a major cause of the difference in sentence structure and expression between Chinese and English. But the importance of punctuation is often overlooked in machine translation. For example, book titles and commas are punctuation marks which are unique to Chinese. During the process of Chinese-English conversion, machine translation will copy them into the target text, resulting in some translation problems.

3.3 Redundancy

Redundancy refers to the functional repetition, overlapping or redundant expressions in the translation (Cui & Li, 2015, p. 21). Repetition is a typical feature of Chinese, and synonymy with different words in the form of four-character words is very common for emphasis. But English avoids repetition and often uses alternatives such as pronouns and prepositions to replace the repeated part. In the process of machine translation, the machine will strive to translate all the content of the source language, which is alien with English expression habits.

3.4 Lexical Vacancy

Lexical vacancy refers to the difficulty in achieving complete equivalence between the source language and the target language, resulting in lexical vacancy in translation. Due to differences in culture and social background between China and other countries, lexical vacancy occurs in the translation process,

45

including cultural defaults during the translation of culture-loaded words. Moreover, in the ever-changing modern society, in addition to new words emerging with the development of the times, many old words have also been given new meanings, such as "韭菜" (it once meant a kind of vegetables, and now also refers to shareholders taking a beating) and "香菇" (it once meant a kind of vegetable, and now also refers to a state of sadness). Therefore, if the lexical vacancy cannot be solved, the translation quality cannot be guaranteed. Because the society is changing so fast that the original database of machine cannot keep up with the pace, or the research on the cultural differences is not sufficient, the machine translation cannot accurately detect the meaning of the source text, which will lead to such translation problems.

4. Translation Strategies

The above-mentioned problems could be solved by adopting translation techniques including replacing the inconsistent terms in source text and adjusting the division of sense group in pre-translation editing, as well as addition, omission and shift in post-translation editing.

4.1 Translation Strategies from the Perspective of Pre-editing

There are differences between Chinese and English languages. Chinese is featured with run-on sentences, four-character words, and implicit subjects. At the same time, it stresses on parataxis with more active sentences. English is featured with complex and long sentences, as well as explicit subjects, and attaches importance to hypotaxis. One of the main tasks of pre-editing Chinese is to eliminate the ambiguity of the source text, that is, to convert human-comprehensible words and sentences into machine-comprehensible ones (Zhong, Xu, & Li, 2021, p. 36). Therefore, this section focuses on pre-translation editing from the aspects of replacing the well-translated terms and modifying punctuation marks so as to solve term inconsistency and incorrect punctuation mark segmentation.

4.1.1 Replacing the Inconsistent Terms in Source Text

Although machine translation has powerful memory and terminology functions, it is not a panacea. Chinese culture has a long history, and neologisms emerge endlessly on the Internet. The computer storage function is not nearly enough to catch up with their development and evolution. Faced with some specific words, especially those with cultural connotations or implied meanings which are not included in the given glossary, machine translation cannot understand and translate correctly. In this case, words in the source language can be replaced in advance through pre-translation. Besides, affected by the context, there will be inconsistent terms in machine translation, which weakens the coherence and cohesion of the translated text. If the expression of the same term is inconsistent, it also will increase the reading difficulty of readers, and may even cause misunderstanding. In view of this, the translation and replacement of inconsistent terms in advance can improve the consistency and the quality of machine-translated text. Moreover, after the terminology is established, it can also provide reference for future texts with similar topics.

By replacing the inconsistent terms in source text, the whole text is unified, so that the language and expression of the machine-translated text are more natural and fluent, which not only improves the quality of the translated text, but also reduces the workload of post-translation editing.

4.1.2 Adjusting the Division of Sense Group

Adjusting the division of sense group refers to various adjustments such as changing, deleting, editing, and omitting the source text, in order to meet the requirements of the "patron" mainstream social ideology (Yi & Liu, 2021, p. 4). Due to the inherent characteristics of Chinese, such as run-on sentences and four-character words, it is inevitable to cause difficulties including structural ambiguity in the process of machine translation. Therefore, it is necessary to analyze the original text and try to adjust the division of sense group that is difficult for machine recognition in the source text, so as to improve machine translation's recognition ability of the source text and improve translation quality.

There are differences between Chinese and English punctuation marks. For example, some punctuation marks in Chinese are not found in English, such as stops and book titles; some punctuation marks in English are not found in Chinese, such as hyphens; most punctuation marks are both used in Chinese and English, but their usages are slightly different. In the process of machine translation, the machine translates the text according to the arrangement and structure of punctuation marks in the original text, resulting in "deficient" or "repetitive" translation. In such cases, modifying punctuation marks and reforming sentence structure can change the source text into a language that is easier to understand and translate by the machine, thus improving the readability of the text.

Punctuation marks are an indispensable part of a written language, mainly used to indicate the nature and function of sentences (Yuwen, 2001, p. 76). It is also a detail that is easy to be ignored in the translation process, but it reflects the translator's ability of avoiding "stupid mistakes" and the level of professional qualification. In addition, the modification of punctuation marks is also an important aspect of pre-translation editing. This is similar to the text alignment function of many computer-aided translation tools such as Trados, but text alignment is to manually segment the machine-translated target language translation and the source language, rather than better modify the source language text to make machine translation more recognizable.

4.2 Translation Strategies from the Perspective of Post-editing

Post-translation editing refers to the process of modifying the source output of machine translation for certain purposes (Feng & Cui, 2016, p. 67). Liu Yi stresses that no matter how intelligent the machine is, and no matter how long the translator spends on pre-translation editing, the texts that are not successfully translated still account for a certain proportion (Liu, 2014, p. 101). At present, for the gap between professional human translation and machine translation, we cannot rely on machine translation alone, thus post-translation editing is also required. No matter the language expressions are familiar or not, it is still difficult to achieve the prefect translation result, so post-translation editing is of great importance (Yang & Fan, 2021, p. 59). Thus, the author puts forward translation skills such as addition, omission and shift to improve the accuracy and quality of machine translation.

4.2.1 Omission

Omission refers to a method of removing redundant information in translation and making the translation more fluent (Wang & Wen, 2020, p. 134). It is not simply to delete the content, but to do some meaningful deletion avoiding repetition and redundancy. The meaning is still complete after deletion, but the expression is more concise. In order to achieve neat writing, Chinese often uses a large number of balanced and parallel structures and a string of four-character sentences. But English is concise and avoids repetition. Therefore, in the process of Chinese-English translation, repetitive problems will inevitably occur, and omission can effectively reduce the redundancy to make the edited language more suitable for English expression habits.

4.2.2 Addition

Addition refers to the appropriate addition to the translated text, so that it not only faithfully expresses the content of the source text, but also conforms to the expression habits of the target language. In Chinese-English translation, addition includes the amplification of pronouns, background knowledge and subjects. The text translated by machine is right sometimes, but the machine fails to fully take into account the knowledge reserves of the target language readers. With supplementary information, it can not only better convey the original information, but also reduce the barriers that appear in cultural exchanges. For example, in the past, Chinese culture may belong to the "weak culture", and the translation of classic books mainly adopts the strategy of "domestication". However, as Chinese culture is increasingly prosperous, translators who are familiar with the similarities and differences of the two languages and cultures (Zhang, 2010, p. 26) should shoulder the responsibility of helping Chinese culture "going global", so that more people can understand the excellent traditional Chinese culture.

In addition, translation skills such as adding footnotes, paraphrasing, or adding direct annotations can make up for the default of the culture-loaded words. Culture-loaded words, also known as culture-specific words, refer to the characteristic words (including idioms and common sayings) that carry a certain cultural connotation and show folk customs, which are often concentrated in national cultural works (Sun & Han, 2021, p. 91). It is this exclusivity that makes it a big obstacle in translation. In the process of translation, rigid translation will damage readability, and the practice, "completely change" will make it lose its original taste. Therefore, the author should strive to reduce cultural differences while enhancing readers' reading experience and cognitive ability.

4.2.3 Shift

Shift refers to the process of converting the language unit or structure of the original language into a language unit or structure with similar properties, or corresponding properties, or heterogeneous properties in the target language (Xiong, 2014, p. 87). Simply put, it is to change the form of expression to a certain extent, but the original meaning remains unchanged. When "hard translation" or "rigid translation" occurs in machine translation, this translation strategy can be adopted. For example, when translating dialects, idioms, and folk songs, the machine cannot understand the cultural connotations in them, but often translate word by word, making the entire text incomprehensible. In view of this, these

translation difficulties are manually converted into expressions that are familiar to the target language readers and easy to understand.

Machine translation has always been a hot topic in translation circles at home and abroad, and has high requirements on the ability of researchers themselves. The author finds that machine translation has difficulties in cultural factors such as beliefs and customs mainly because it fails to understand the cultural context and emotions conveyed in the text. The biggest difference between human and machine is the feelings. The translators will consider the differences and taboos of different cultures and the reaction and mood of the target language readers, rather than simply complete the translation task. Therefore, in order to promote the successful exchange between Chinese and foreign cultures, post-translation editing is also indispensable.

5. Conclusion

With the progress of science and technology and the continuous development of artificial intelligence, machine translation has attracted more attention. It is faster and cheaper than human translation, and can handle huge volumes of multilingual text. However, machine translation lacks logical thinking, which will make the translation lack of logic and standardization, thus leading to ambiguity, or even distortion of the source text. The "pre-editing + machine translation + post-editing" model gives full play to the advantages of the interactive cooperation between human and machine by replacing the inconsistent terms in source text and adjusting the division of sense group in pre-editing part which improves the translatability of the text, and increasing the readability of the translation by addition, omission and shift in post-editing, both greatly improve the translation quality and have strong operability.

References

Cui, Q. L., & Li, W. (2015). The Character of Error Types of Post-editing: Perspective of Machine Translation Based on Scientific and Technology Materials. *Chinese Science & Technology Translators Journal*, *28*(4), 19-22.

Cui, Q. L. (2014). On Post-editing in Machine Translation. *Chinese Translators Journal*, *35*(6), 68-73.

Feng, Q. G., & Cui, Q. L. (2016). Research Focuses and Trends in Post-editing of Machine Translation. *Shanghai Journal of Translators*, *6*, 67-74+89+94.

Feng, Q. G., & Gao, L. (2017). On the Influence of Pre-editing in Accordance with Controlled Language on Machine Translation. *Contemporary Foreign Language Studies*, *2*, 63-68+87+110.

Gao, L. L., & Zhao, W. (2020). An Overall Study on Machine Translation. *Foreign Languages in China*, *17*(6), 97-103.

Guo, G. P., & Wang, Z. Y. (2017). Research on the Pre-edit and Post-edit of Machine Translation in Science and Technology Text Translation. *Journal of Zhejiang International Studies University*, *3*, 76-83.

Liu, Y. (2017). Recent Advances in Neural Machine Translation. *Journal of Computer Research and Development*, *54*(6), 1144-1149.

Liu, Y. (2014). Empirical Study of Controlled Language for Improving the Quality of the Machine Translation Technical Specifications. *Journal of Chongqing University of Technology (Natural Science)*, *28*(3), 96-101.

Mellinger, C. (2017). Translators and Machine Translation: Knowledge and Skills Gaps in Translator Pedagogy. *The Interpreter and Translator Trainer*, *4*, 280-293. https://doi.org/10.1080/1750399X.2017.1359760

O'Brien, S. (2006). Methodologies for Measuring the Correlations Between Post-Editing Effort and Machine Translatability. *Machine Translation*, *1*, 37-58. https://doi.org/10.1007/s10590-005-2467-1

Sun, L., & Han, C. H. (2021). English Translation of Culture-loaded Words in *Folding Beijing* from the Perspective of Eco-translatology. *Shanghai Journal of Translators*, *4*, 90-94.

Wang, N., & Wen, Y. Y. (2020). On the Translation Problems of Luo Xiwen Version of *Shanghanlun*. *Chinese Translators Journal*, *41*(5), 130-135.

Xiong, B. (2014). Concept Confusion in Translation Studies: A Case study of "Translation strategies", "Translation methods" and "Translation techniques". *Chinese Translators Journal*, *35*(3), 82-88.

Yang, W. D., & Fan, Z. R. (2021). Case Studies of Postediting in Machine Translation of Scientific and Technological Texts. *Shanghai Journal of Translators*, *6*, 54-59.

Yi, X., & Liu, J. P. (2021). On Mistranslation, Rewriting, and Omission in the English Translation of *The Sky Dwellers*. *Foreign Language and Translation*, *28*(1), 1-8+98.

Yuwen, G. F. (2001). The Difference and Proper Usage of Punctuation Marks in Chinese and English. *Chinese Journal of Scientific and Technical Periodicals*, *12*(1), 73-76.

Zhang, Y. (2010). The Loss and Gain of Interpretation: A Case Study of Amplification in the English Translation of *Gan Xiao Liu Ji*. *Contemporary Foreign Language Studies*, *4*, 26-29+62.

Zhong, W. M., Xu, J. M., & Li, Y. X. (2021). Pre-editing Mechanisms for Vagueness of Chinese Sci-tech Text in Machine Translation. *Chinese Science & Technology Translators Journal*, *34*(4), 36-39+42.