On the Strategies of Machine Translation Post-Editing Based on Cloud Platform

Lihua Qing

School of English Language and Culture South China Business College, Guangdong University of Foreign Studies Guanghzou, China

Email: 65993576@qq.com

Abstract. In the era of AI, machine translation post-editing(MTPE) based on cloud platform is an important part in the translator's translation environment. This study analyzes the existing problems in machine translation, and puts forward specific post-editing strategies from the aspects of lexis, syntax, text, style and format, in order to provide some references to improve the quality of post-editing based on computer-aided translation software.

Keywords: cloud platform, machine translation, post-editing, strategies

1. Introduction

In the era of AI, cloud platform is a computer technology that develops very rapidly in recent years, and has made great progress in many application fields. [1] The cloud platform can be used to solve the contradictions between calculating and storing resources and implementation costs in a single network terminal, in which multiple language translation tools or systems can be deployed. Major technology and Internet technology companies have launched Google Translate, Baidu Translate, Microsoft's Bing Translate, Youdao Translate and so on, which can be used as software plugins in the computer-aided translation (CAT) software as the cloud platform.

Post-editing or postediting (PE for short) refers to the process of processing and revising the source text of machine translation according to a specific purpose, including correcting translation (language) errors and improving the accuracy and readability of machine translation. MTPE refers to machine translation post-editing. MTPE is an important approach to improve the quality of machine translation and has become a heated topic in industry application and academic research. In recent years, scholars at home and abroad have made some researches to explore the quality of MTPE. Koponen and Salmi designed post-editing experiments to explore how subjects could improve their post-editing translations and whether post-editing could improve the efficiency and quality of translation.[2] Nunes-Vieira predicted machine translation errors, pointed out MTPE processes, and determines the feasibility of MTPE. [3] Zhou Hong, Zhou Mingli and Jiang Sijia analyzed the difficulty of cloud platform technology in the application process and proposed feasible application modes based on a large number of corpus translation practices. [4] Zhao Qiurong and Fu Jiying discussed post-editing strategies of non-subject sentences in government official documents based on Google-translated text as post-edited text.[5] This study analyzed the existing problems in machine translation, and proposed specific post-editing strategies from the perspective of lexis, syntax, text, style and format, in order to provide some references to improve the quality of MTPE based on CAT software.

2. The existing problems in machine translation

Due to the complexity of the language, the translation quality of the cloud platform is poor, and there exist many errors or improper translation, which has become the main task of the current post-editing work. Common errors consist of the ones in lexis, syntax, text, style and format.

2.1 Errors at the Lexical Level

There are some lexical errors in machine translation, including misspelling errors, number errors, terminology errors, acronym errors, multiple translation or untranslated expression and over-translation. Vocabulary is the critical part in translation. Lexical errors directly influence the sentence comprehension. Basically, machine translation takes literal translation. For example, The Chinese word "高压" is sometimes translated into "high pressure" and sometimes into "high voltage" by the cloud platform. "水砂纸"is translated into "water sandpaper" rather than "waterproof sandpaper". The final choice depends on the post-editing translator.

2.2 Errors at the Syntax Level

Chinese emphasizes parataxis while English highlight hypotaxis. In the machine translation, sentence mistranslation frequently appear, such as part of speech, word order, conjunction, phrase, voice, infinitive, participle and clause. Besides, there are element absence and structural misuse. The lexical error influences the understanding of the discourse. Machine translation is sentence by sentence, and there is no connection or correspondence between the contexts, so the translation will be very boring and verbose, which is the so-called machine translation trace. Therefore, it is necessary to add some conjunctions, such as but, and, however and so on.

2.3 Errors at the Discourse Level

In the level of discourse, the mistranslation phenomenon like redundancy, illogical content and coherence exists. Discourse errors are a big problem in translation. The cloud platform can not distinguish the logic in the whole text, which cause many errors in the target text, The errors increase the reading burden for the post-editing translators.

2.4 Errors at the Style Level

According to the writing style of the source text, the post-editing translator should analyze the writing style of the translated text, including argumentation, narration, exposition and description. After determining the writing style, the post-editing editor need to keep the style of the target text consistent. In addition, the language styles also need to be considered. Thus, it is quite essential to keep the consistency in post-editing.

2.5 Errors at the Format Level

In terms of format, there are punctuation errors, improper units, and case issues. In   machine translation, it is necessary to abide by the format. Machine translation cannot accurately recognize the format errors in the translation.

3. Strategies of mtpe

Generally speaking, there are different types of post-editing according to the degree of manual intervention, which can be divided into light post-editing and full post-editing. The former refers to unnecessity to translate a text comparable to artificial translation in order to obtain a barely comprehensible text. Full post-editing is to obtain a target text with high quality that can be equivalent to that of artificial translation. [6] The following principles and strategies of post- editing are put forward from the perspective of full post-editing.

3.1 Principles of MTPE

First, it is critical to achieve accuracy in meaning and style. On the one hand, the post-editing translators should accurately understand the original content and convey it correctly by using CAT translation functions. We cannot add or delete the original meaning while revising, keeping in line with the original content and form. On the other hand, there are differences in terms of wording,

tone and format between different genres, which requires post-editing translators to be considerate. They should not only have a solid foundation in literature, cross-cultural communication literacy, master rich knowledge in nature science or humanities, but also understand the styles differences between English and Chinese.

Secondly, the post-editing translator should identify all the errors, aiming to revise the errors. Post-editors need to master a large amount of grammar knowledge, summarize the common types of errors through a lot of practice, so as to achieve faster and more accurate recognition of errors. Therefore, it is necessary to continuously improve the competence of post-editing, including proficient employment of source language and target language, software application competence, good word processing and editing competence, etc.

Thirdly, combination between man and machine can reach the double results with half effort. Machine translation is time-consuming and efficient, but the quality of machine translation is not good. In contrast, artificial translation takes a longer time, but the translation quality and credibility are excellent. In post-translation editing, error correction functions of other software can be applied to improving the accuracy rate.

3.2 Methods of MTPE

In the first place, it is critical to create a termbase(TB), In view of lexical errors, post-editors should establish a TB of professional terms, a TB of acronyms and a TB of high-frequency conjunctions, constantly expanding the database capacity and improve the correct matching degree. In addition, the post-editor should choose the correct term expression, select the specific meaning according to the context, and refer to authoritative websites such as UNTERM and Linguee.

In the second place, The post-editing translator should take full advantage of syntax retrieval tools. For syntactic errors, the post-editing translator can first use a grammar retrieval tool, such as Grammarly, to identify syntactic errors, and then can analyze and modify the syntactic errors correspondingly.

Thirdly, in terms of discourse errors, it is necessary to divide and restructure the sentences. In order to solve the problems of missing translation, multiple translation, sentence breaking and missing elements in a text, the post editor should split and reorganize the sentence structure and adjust the word order reasonably. At the same time, it is appropriate to highlight the theme information, use means such as repetition, omission and connection, pay attention to the appropriateness of the context, and reorganize sentences and paragraphs if necessary.

Last but not least, it is also important to standardize the language styles and the format. In accordance with the source text, the post-editor should follow the basic format norms and keep the consistency of language styles.

4. Conclusion

Post-editing competence is closely related to translation competence, which can be called translation competence plus. They need to master professional knowledge and post-editing skills. Just like translation competence, the post-editors need professional training and a lot of practice to be qualified for MTPE. Post-editing work, especially at the beginning, requires expectations and preparation. Instead of blaming and pursuing perfection, they should be no feeling of loss or resistance, but rather a discerning eye to find more usable content from the machine translation. Post-editing translators can improve the translation quality of machine translation and meet the requirements of full post-editing by adopting targeted post-editing strategies for different types of errors in machine translation.

Acknowledgment

References

[1] L. Yupeng, Q. Mingxiu, Z. Shilei, et al. Deep Fusion of Large-scale Features in Statistical Machine Translation . Journal of Zhejiang University (Engineering Science), 2017 (1) , pp.46-54

[2] M. Kooponen and L. Salmi. Post-editing Quality: Analysing the Correctness and Necessity of Post-editor Corrections[J]. Linguistica Antverpiensia. New Series: Themes in Translation Studies.2017,16, pp.137-148

[3] L. Nunes-Vieira. Cognitive Effort and Different Task Foci in Post-editing of Machine Translation:A Think-aloud Study[J]. Across Languages and Cultures.2017,18(1), pp.79-105

[4] Z. Hong, Z. Mingli, J. Sijia. Research and Practice of Computer-aided Translation Based on Cloud Platform, Popular Science and Technology [J]. 2021(258), pp.10-12

[5] Z. Qiurong, F. Jia. Post-editing Based on Google Translation[J]. Metallurgical Education in China.2018(4), pp.59-63

[6] Z. Tao. Current Situation and Problems of Post-editing in Machine Translation [J]. Foreign Language Teaching, 2021 (4) ,pp.100-104