

Machine Translation and Post-editing in Foreign Language Teaching and Learning: A Systematic Review

Xu Chen

Foreign Language School, Linyi University, Linyi, 276000, China

Keywords: Machine Translation (MT), Post-editing (PE), Post-editing Effort (PEE), Training

Abstract: In recent years, the improvement of machine translation (MT) has facilitated advancements in the new translation mode of Machine Translation Post-editing (MTPE). This study compiles and synthesizes existing literature on post-editing, specifically focusing on text type, evaluation, and pedagogical implication. It is conducted by Systematic Reviews and Meta-Analyses (PRISMA) framework. The findings suggest that the majority of source texts utilized for MTPE in language learning involve in the general domain, with newspapers being the most frequently employed option. Meanwhile, the results indicate that the translation outputs of NMT post-editing and from-scratch translation were comparable when evaluating PE outcomes in terms of accuracy and fluency. Furthermore, the framework proposed by Krings [1] for assessing Post-editing Effort (PEE) has gained significant acceptance in the field. Ultimately, the utilization of NMT can yield benefits and efficacy in the field of language acquisition. The review also suggests future research directions to analyze issues and advance regarding to post-editing.

1. Introduction

The worlds of translation and multilingualism have changed in 2016 as a result of a new method named machine translation (MT) with the development and distribution of artificial intelligence, big data, and neuroscience research results. The combination of these novel techniques gave rise to Neural Machine Translation (NMT). Consequently, post-editing (PE) is now a required talent for qualified translators. In light of these remarkable advancements, systematic research on post-editing is still required. The research questions that were formulated to explore the most vital topics are as follows:

- 1) Which text types and genres were mostly used for machine translation and post-editing (MTPE) in language learning?
- 2) When evaluating MTPE with metrics for translation quality assessment (e.g., with dimensions such as fluency, accuracy, etc.), in which dimension(s) can MTPE show significant improvements?
- 3) How was post-editing effort (PEE) evaluated?
- 4) What's the pedagogical implications of using MTPE in language learning?

2. Methods

This synthesis will be conducted by PRISMA framework. The workflow consists of three steps: identification, screening, and inclusion. In the identification phase, the Scopus database and the SSCI journal *Language Learning & Technology* were searched for conducting this review. The period begins from 1 January 2016 to 1 May 2023, on account of the year 2016 as the launch year of neural machine translation, which enhanced the quality and readability of MT. There were 650 results to be identified by inclusion and exclusion criteria (retrieved on 28, May, 2023).

The following are the inclusion criteria that must be met for eligibility:

- 1) Publications released between 2016 and 2023.
- 2) Researches findings disseminated through a scholarly publication that has undergone rigorous evaluation by experts in the field.
- 3) Researches dealing with the genres/quality/evaluation/impact of MT and post-editing on language learning.
- 4) Researches conducted through empirical or case study methods.
- 5) Academic literature composed in the English language.

In the screening and inclusion phase, the collected articles received a process of eligibility assessment. The process of screening for relevant articles involved analyzing the titles, keywords, and abstracts of 158 papers. Subsequently, 72 full-text versions of papers were identified for verification in the following inclusion phase. Finally, there were 9 articles to be included in this systematic review.

3. Results

3.1 Text Types and Genres

The texts utilized for Machine Translation Post-Editing (MTPE) in the context of language acquisition predominantly belong to the general domain, with newspapers being the most frequently employed option. The rationale behind this is that the text ought to be concise enough to be completed by participants within a single session and should not necessitate any specialized prior knowledge. One study [2] has been conducted utilizing domain-specific texts, specifically manuals relating to medicine and advertisement.

3.2 Translation Quality Assessment

According to Jia et al. [2], the assessment outcomes indicate that the fluency and accuracy of the translation outputs from NMT post-editing and from-scratch translation were similar for both general language texts and domain-specific texts. The study conducted by Zhang and Torres-Hostench [3] revealed that the implementation of MTPE training did not result in a statistically significant improvement in accuracy for students who engaged in post-editing. The error of accuracy is characterized by its openness, making it challenging to identify as it relates to the intended meaning of a sentence rather than a specific case of word misuse. To assist students in identifying and correcting accuracy errors, more and better instructional resources must be generated. Jia and Lai [4] discovered that when translating metaphorical expressions from English to Chinese, the overall number of errors in fluency was significantly lower than that in accuracy for both the HT and PE, regardless of task types.

3.3 Post-editing Effort (PEE) Assessment

The majority of studies that assess the process of physical education utilize Krings' [1] theoretical framework as a foundation. Krings [1] delineates three interrelated dimensions of post-editing effort (PEE) - temporal, technical, and cognitive - in a comprehensive early investigation on post-editing. Temporal effort refers to the total duration of time necessary to complete a given task. Technical effort refers to the mechanical or physical work required during post-editing activities. Cognitive effort refers to the level of mental processing required for the evaluation of post-editing decisions.

The most frequently utilized indications in the assessment process are keystrokes, fixation, pause density, and duration. This helps to enhance comprehension of the evaluation process and to better describe how PE works. In addition, Yamada [5] utilized the concept of perceived ease to assess cognitive effort. The study's participants were instructed to quantify the amount of effort they spent on the translation task by assigning it a numerical value (set at 100) that corresponded to their perception.

3.4 Pedagogical Implication

According to Yamada's [5] research, post-editing (PE) demands a level of proficiency comparable to that of human translation (HT), as neural machine translation (NMT) generates errors that closely resemble those made by humans, making them more challenging for students to correct. According to Jia et al.[4], the research revealed that students encountered difficulties in post-editing, which required specialised translation skills and strategies. According to Chung's [6] declaration, PE practice can serve as a suitable means of assessing learners' underlying linguistic knowledge and fostering their linguistic awareness of the distinctions and similarities between their L1 and L2. This view is corroborated by the findings of Zhang and Torres-Hostench [3].

According to Yang and Wang [7], the enhancement of students' MTPE competence necessitates the consideration of self-regulatory ability as a fundamental element. The study examined the roles and interplay of self-regulation, critical thinking, and motivation in the context of MTPE. The results indicated that self-regulation plays a crucial role in forecasting the performance of MTPE. Hence, it is imperative to consider incorporating self-regulatory teaching activities to enhance the efficacy of training.

4. Discussion & Conclusion

This review aims to thoroughly investigate post-editing in language education. For the PE project's decision-making, text type considerations are crucial. Extremely limited, highly standardized, repetitious, and less imaginative texts are well-suited for machine translation. This may help to explain why news and technical texts are frequently used as the source text in included research designs.

Concerning quality assessment, the results indicate that the translation outputs of NMT post-editing and from-scratch translation were comparable. However, in terms of accuracy, the results were inconsistent. When it comes to process assessment utilizing PEE, the most frequently utilized indications in the reviewed studies are keystrokes, fixation, pause density, and duration.

Additionally, the increasing use of the PE approach raises concerns regarding instruction and education. Attention should be paid to the difference between PE and human translation. In reviewed studies and other related studies [8] [9], this view has reached a consensus: NMT can be advantageous and effective for language learning, but instructors must provide their students with appropriate guidance and training on its use, particularly post-editing practices.

Overall, some implications could be found for future research. Teachers must be aware of their

students' second language (L2) proficiency in order to make effective use of NMT. Despite a general consensus on the value of PE training, this review's findings highlight a lack of data about the application of NMT with specific training in the classroom.

References

- [1] Krings, H. P. *Repairing Texts: Empirical Investigations of Machine Translation Post-editing Processes*, vol. 5, 2001. Ohio: Kent State University Press.
- [2] Jia, Y., Carl, M., & Wang, X. *How Does the Post-Editing of Neural Machine Translation Compare with From-Scratch Translation: A Product and Process Study*. *The Journal of Specialised Translation*, vol. 1, no. 31, pp. 60–86, 2019a.
- [3] Zhang, H., & Torres-Hostench, O. *Training in Machine Translation Post-editing for Foreign Language Students*. *Language Learning & Technology*, vol. 26, no. 1, pp. 1-17, 2022.
- [4] Jia, Y., & Lai, S. *Post-Editing Metaphorical Expressions: Productivity, Quality, and Strategies*. *Journal of Foreign Languages and Cultures*, vol. 6, no. 2, pp. 28-43, 2022.
- [5] Yamada, M. *The Impact of Google Neural Machine Translation on Post-editing by Student Translators*. *The Journal of Specialised Translation*, vol. 31, no. 1, pp. 87-106, 2019.
- [6] Chung, E. S. *The Effect of L2 Proficiency on Post-editing Machine Translated Texts*. *The Journal of Asia TEFL*, vol. 17, no. 1, pp. 182–193, 2020.
- [7] Yang, Y., & Wang, X. *Predicting Student Translators' Performance in Machine Translation Post-editing: Interplay of Self-regulation, Critical Thinking, and Motivation*. *Interactive Learning Environments*, vol. 31, no. 1, pp. 340–354, 2020.
- [8] Chung, E. S., & Ahn, S. *The Effect of Using Machine Translation on Linguistic Features in L2 Writing across Proficiency Levels and Text Genres*. *Computer Assisted Language Learning*, vol. 35, no. 9, pp. 1–26, 2021.
- [9] Koponen, M., Salmi, L., & Nikulin, M. *A Product and Process Analysis of Post-Editor Corrections on Neural, Statistical and Rule-based Machine Translation Output*. *Machine Translation*, vol. 33, no. 1-2, pp. 61-90, 2019.