ISSN 1799-2591
Theory and Practice in Language Studies, Vol. 14, No. 7, pp. 2025-2036, July 2024
DOI: https://doi.org/10.17507/tpls.1407.09

Measuring the Efficiency of Post-Edited Text Generated by CAT Tools: An Experimental Study

Hind S. Alsaif
Decision Support Centre, Royal Court, Riyadh, Saudi Arabia

Ebtisam S. Aluthman*
Department of Applied Linguistics, College of Languages, Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia

***Abstract*—Motivated by the technological advancements in computer-assisted translation (CAT) tools and the notable lack of academic research regarding their application in the Arabic-translation context, this study aims to investigate the differences in translators' performance when comparing traditional human translation and post-edited CAT tool-generated text in terms of speed and effort. This study investigates the performance of professional translators in Saudi Arabia through traditional translation from scratch (TFS) and post-editing (PE) approaches. Data was collected from nine translators with 5–12 years of experience who had exposure to CAT tools. The participants translated an Arabic educational article into English using both methods. This study utilized Phrase CAT and Translog-II software to analyze the participants' time and keystrokes. The results indicate that PE was significantly faster than TFS, with PE requiring 65.1% less time. PE also demanded significantly fewer keystrokes, suggesting lower technical effort. Correlations between keystrokes and time indicate a strong positive relationship in PE, implying that more technical effort correlates with increased temporal effort. These findings emphasize the efficiency of PE in enhancing productivity and suggest the importance of CAT tools and PE training for translators to meet industry demands effectively. Furthermore, this study underscores the need for continuous updates in CAT tool courses and the integration of PE training to prepare translators for constantly evolving technological landscapes.**

***Index Terms*—post-editing, traditional human translation, translation technology, computer-assisted translation (CAT) tools, translation memories**

I. Introduction

The development of interactive translation technologies began in the late 1970s and early 1980s, a period marked by significant advancements in the field (Kay, 1980; Melby, 1979, 1981, 1982). This era marked a paradigm shift in translation studies, with researchers exploring various methodologies for leveraging tools and computers to enhance translators' work (Hargrave & Savourel, 1997). This exploratory phase led to the establishment of what is now known as computer-assisted translation (CAT), a term coined and elaborated upon by Hutchins (1998). According to Bowker (2002), CAT technology includes a broad range of computerized tools that aid translators in their professional tasks: "CAT technology can be understood to include any type of computerized tool that translators use to help them do their job" (p. 6). This definition highlights a distinct difference between CAT tools and machine translation (MT). When using CAT tools, a human translator performs the translation, but computer assistance is incorporated into the process (Christensen & Schjoldager, 2010).

This human role of translators has significantly contributed to the growing practice of post-editing (PE), in which translators or post-editors refine machine-generated translations to meet specific quality standards. PE has emerged as an essential process, capitalizing on the strengths of MT while leveraging human expertise to ensure accurate and fluent translations. In the early phases of the editing process, a post-editor selects the correct translation and rearranges it according to the rules of the target language based on suggestions provided by a computer dictionary (Hutchins, 1986; as cited in Koponen, 2016a). According to Hutchins (1986, p. 73), the purpose of PE is "to produce out of the raw output […] a readable translation in a fraction of the time it would take a bilingual expert to produce a translation with the traditional procedure".

The aforementioned rationale has prompted this study, which involves an experiment aimed at assessing the performance of professional translators in both PE and traditional translation approaches. The main objective is to contribute to the advancement of translation technology research and to direct the translation industry's attention toward

* Corresponding Author. Email: esaluthman@pnu.edu.sa

the impact of translation technology, especially within the Arabic-language context. Nine professional translators from Saudi Arabia were chosen to participate in the experiment, and their performance was assessed in terms of speed (time taken) and effort (number of keystrokes). In the PE approach, the translation was generated using the Phrase CAT tool, which is known for its high-quality features. Translog-II software was utilized to record the time spent by the translators on each approach and the number of keystrokes they made. This research is expected to contribute to the existing literature on the assessment of the efficiency of PE technologies, especially within the Saudi translator community, where there is a lack of studies on this topic. This study aims to address this gap and investigate the differences in translators' performance via a comparison between traditional human translation and post-edited CAT tool-generated text in terms of speed and effort. Accordingly, the research questions of this study are as follows:

1. RQ1: Which of the two approaches (PE of CAT tool-generated text and traditional human translation) requires more time (speed)?
2. RQ2: Which of the two approaches (PE of CAT tool-generated text and traditional human translation) requires more keystrokes (effort)?

Furthermore, this research is designed to test two primary hypotheses:

Hypothesis 1: Performing PE on CAT tool-generated text helps with expediting the process, which results in higher productivity rates.

The study aims to determine whether the process of performing PE on CAT tool-generated texts leads to a quicker translation process compared to the traditional method of human translators translating texts entirely from scratch (translation from scratch [TFS]). By measuring and comparing the time taken to complete translations using these two approaches, this research intends to provide insights into the efficiency and practicality of PE with CAT tools relative to conventional human translation methods.

Hypothesis 2: PE is less cognitively demanding than traditional human translation.

Testing the validity of this hypothesis entails comparing the effort involved in the two aforementioned translation methods. In this context, "effort" is quantified by the number of keystrokes needed to complete a translation. A higher number of keystrokes typically signify more input from the translator, suggesting greater effort. By analyzing and contrasting the keystroke counts for both post-edited CAT tool-generated texts and traditional human translations, this research provides insights into the relative effort demanded by each method. This comparison is crucial for understanding the efficiency and workload implications associated with the use of CAT tools in translation processes.

II. PE IN TRANSLATION

A. Definition and Classification

Veale and Way (1997; as cited in Allen, 2003, p. 297) initially defined PE as the correction of machine-translated text by human linguists or editors, emphasizing the human role in refining machine-generated translations. Allen (2001) expanded on this, describing PE as an integrated process within MT in which professional translators correct machine-generated translations and fuzzy matches from translation memories (TM). This collaborative process aims to produce higher-quality translations in less time. Schäfer (2003, p. 3) further defined PE as the task of refining machine-translated text to achieve an acceptable level of quality for end users, highlighting the focus on enhancing raw MT output for user satisfaction.

PE is classified into two types. These classifications are also known as levels or degrees of PE, which are based on the number of corrections and the effort needed to reach the required quality translation. A few terms have emerged to describe these classifications, such as light and full PE; rapid and conventional PE; and partial and full PE (Arenas, 2019). In the 1990s, rapid PE was commonly referred to as "minimal or minimum" in the industrial and corporate sectors (Allen, 2003). In light PE, the translator lightly edits the MT translation to produce "understandable and usable" output (Hu & Cadwell, 2016), as the main purpose of the translation is to be understandable and grammatically correct. Allen (2003) explained the concept of light PE as a process in which "the post-editor ensures there are no linguistic or terminology errors or mistranslation. Its main purpose is to make the text understandable without altering its style" (p. 297). In contrast, a full PE process emphasizes producing a translation that is human-like and sounds as natural as possible (Hu & Cadwell, 2016). Similar to light PE, full PE should also eliminate any grammatical errors, typographical errors, punctuation errors, and other basic issues. Furthermore, it should address stylistic aspects, bearing native readers in mind (Allen, 2003). In essence, the post-editor should also pay attention to the cultural references of the translated text to make it as fluent as possible (Densmer, 2014).

B. The Speed of PE and Traditional Human Translation

The first hypothesis of this study posits that PE may be a more practical method than TFS, potentially leading to increased translation productivity. This hypothesis is grounded in the notion that CAT tool-generated text, which involves refining and correcting machine-translated texts, could streamline the translation process. The rationale behind this hypothesis is that MT provides a base translation that can be modified and improved by a human translator, potentially reducing the time required compared to TFS.

Different measurements of speed have been utilized in empirical studies comparing the speed of PE and traditional human translation. Vazquez et al. (2013) examined the speed of the PE process compared to human translation. They utilized the ACCEPT PE tool to record the time spent editing each section, and it was found that performing PE on MT was the fastest process, while TFS scored lowest in terms of speed. Läubli et al. (2013) proposed a translation productivity experiment to compare PE with computer-aided translation. Their study involved six participants majoring in German-to-French translation, who were allowed to use terminology and translation memories via a fully functional translation workbench. Screen recordings were used to precisely capture time without requiring the participants to track or report it. The controlled experiment demonstrated that PE yields faster translations of consistent quality compared to computer-aided translation. Aranberri et al. (2014) compared the speed of PE among professional translators and untrained users. The participants in both groups were tasked with translating and performing PE on two different texts using the Bologna Translation Service, a tool equipped with statistical MT and time-tracking capabilities. Despite varying factors, such as text complexity and MT quality, the results indicated a productivity gain for PE, with translators' and users' speeds increasing by 17.66% and 12.43%, respectively.

Within the context of Arabic–English translation, Haji Sismat (2016) conducted an experiment involving non-native student translators majoring in Arabic and competent in English. The students, inexperienced with CAT tools but familiar with MT engines, translated legal and journalistic texts and undertook PE tasks using texts pre-translated by Google and Bing Translate, as well as the MemoQ TM CAT tool. Comparing the approaches showed that the participants completed PE tasks faster than TFS, indicating that PE improved the students' speed to levels comparable to those of professional translators. Additionally, Samman's (2022) study explored the productivity of PE and translation tasks, employing paired sample t-tests and repeated measures analysis of variance to compare the total task duration between the control and experimental groups. The results suggested significant productivity gains in both groups, with the experimental group exhibiting higher gains. The comparison between human translation and MT PE methods showed that MT PE led to greater productivity gains. These findings underscore the importance of considering diverse factors and methodologies when evaluating productivity in translation and PE tasks.

C. Translators' Effort in PE and Traditional Human Translation

The second hypothesis of this study is that PE, in addition to having quality equal to that of traditional human translation while being faster, consumes less cognitive effort. O'Brien (2007) conducted a study on PE effort in machine-translated texts to quantify the level of temporal and technical effort. The study included two groups of English–German translators. One group was instructed to post-edit a machine-translated text, while the other was instructed to translate the same text from English into German. Furthermore, text segments containing negative translatability indicators were isolated to demonstrate the difference in effort required to post-edit and translate segments with and without such indicators. Across all segments, PE effort was lower than translation effort.

Vazquez et al. (2013) utilized the ACCEPT PE tool to count keystrokes during each translation process. This method was employed as a metric for measuring effort. Their findings revealed that performing PE on MT involved the least number of keystrokes, indicating lower effort compared to other methods. In contrast, the process of TFS, in which no translation proposals and no MT options were offered, required the highest number of keystrokes, signifying a greater level of effort.

Expanding this investigation, Daems et al. (2017) compared PE and human translation processes among master's students (10 participants) and professional translators (13 participants). The study focused on the effort, among other variables, involved in these translation processes. Articles from newspapers, assessed for complexity and comprehensibility using Lexile measures, served as the texts for translation. The MT inputs were sourced from Google Translate. The CASMACAT PE tool, which can track both keystrokes and eye gaze, was employed to evaluate cognitive effort. The study concluded that the cognitive demand of PE was less for both student and professional groups, with their performance levels being nearly identical.

Collectively, these empirical investigations emphasize that performing PE on TM or MT is more productive than traditional translation methods. The body of research, including studies by O'Brien (2007), Vazquez et al. (2013), and Daems et al. (2017), consistently shows that performing PE on TM or MT demands less effort and yields better quality outcomes than traditional translation. However, it is critical to note the diversity in participant profiles across these studies. Not all participants were professional translators; some were lay users (Aranberri et al., 2014), non-native language students (Haji Sismat, 2016), or translation trainees (Garcia, 2010).

Motivated by the technological advancements in CAT tools and the lack of academic research regarding their application in the Arabic-translation context, the present study aims to address this gap in the existing literature by exploring the utilization of CAT tools among Saudi professional translators. The focus on Saudi translators specifically aims to provide a well-rounded view that enriches the broader understanding of the translation industry. Despite the wealth of literature available, there remains a notable absence of experimental research examining the aspects of effort and speed in English–Arabic translation, a domain yet to be thoroughly explored. The outcomes of this study are expected to be beneficial to translators, academics, and educators, aiding them in recognizing implications, insights, and recommendations that are pertinent to the various facets and sectors of translation at a local level.

III. Materials and Methods

The hypotheses of this study were examined based on the performance of professional translators. Data were collected from nine Saudi translators who possessed bachelor's or master's degrees in translation and who had adequate training in CAT tools. With the exception of one translator, who had 2 years of experience, all translators had 5–12 years of experience working as full-time professional translators. The TM tools reported to be used by the translators were Trados (three translators), MemoQ (one translator), Matecat (three translators), Memsource (one translator), and Wordfast (one translator). Based on the translators' qualifications and experience, nearly all translators had been exposed to and were familiar with both traditional translation and PE approaches. However, from the provided information, the degree of exposure varied among the participants.

The participants were requested to translate the selected texts using both approaches. For this study, an Arabic article (from Mawhiba) was selected for translation into English. This text type was chosen because "the style of a non-literary text generally contains fewer or more controlled ambiguities, gaps, and possibilities for engagement" (Boase-Beier, 2011, p. 76). The selected text, being from an educational genre, did not contain any implied meanings. This choice was also based on the recommendations of three researchers and academics in the field. The researchers involved in the present study reached an agreement that approximately 350 words would be sufficient to indicate the results. Furthermore, the texts were taken from the same article to ensure a parallel level of complexity. Consequently, the final corpus for the study consisted of an Arabic text (330 words) to be translated into English and an English text pre-translated using the Phrase CAT tool (346 words) to be post-edited.

To address the research questions posed in this study, the researchers chose to use two tools for collecting, recording, and analyzing data. The first tool was Phrase CAT, formerly known as Memsource, which was used to produce the text that would be post-edited by the participants and to analyze the quality of their final output. The other tool was Translog-II, which was used to record the participants' time and keystrokes in both approaches. The Translog software was developed in Copenhagen Business School in 1999 under Professor Arnt Jakobsen (Munday, 2016). This software was chosen because its primary function is to collect quantitative data for translation research. Translog-II, the upgraded version of the software, was used in the present study. The program allows the researcher to follow the translation process, recording the particular actions performed (keystrokes) and the amount of time a translator spends on speed, pauses, and arrow movement (Karwacka, 2013). An enormous body of literature has demonstrated the effectiveness of these tools by evaluating the PE approach (Balling et al., 2014; Guerberof, 2009). The use of CAT tools has been perceived as having a fundamental impact on translation accuracy and could transform the translation industry. A screenshot of Phrase is given in Figure 1.

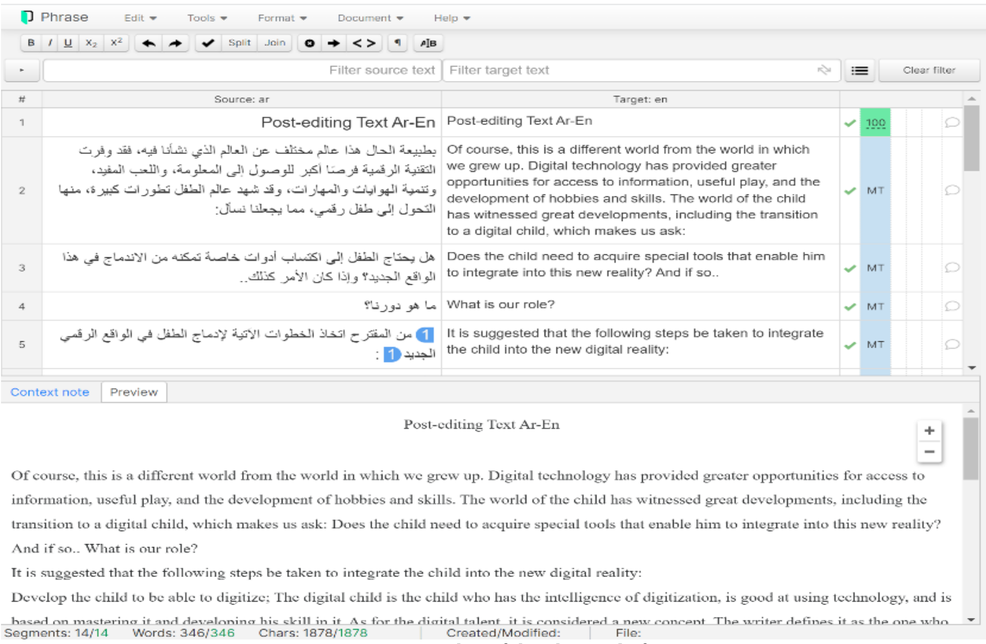

Figure 1. Screenshot of the Phrase Platform

The participants were requested to participate in a language lab, where they used devices to record the process of the PE and TFS approaches on Translog-II. For TFS, the participants were given the text and requested to translate it into English using Translog-II; they were allowed access to dictionaries and online resources but were instructed to avoid using MT. For PE, the participants were given the unedited text and requested to edit all errors in the translation until its quality matched that of human-edited text.

The researchers utilized Translog-II to extract the speed and effort of each participant. The researchers were able to calculate and compare the participants' time and speed because the data .xml file on Translog-II displays the precise start time and end time for both approaches (see Figure 2). To analyze time, the *p*-value was calculated to indicate the statistical significance in relation to the total time difference between the two approaches, as the *p*-value could serve as an indicator of whether the observed differences resulted from randomness or reflected the actual differences between the approaches.

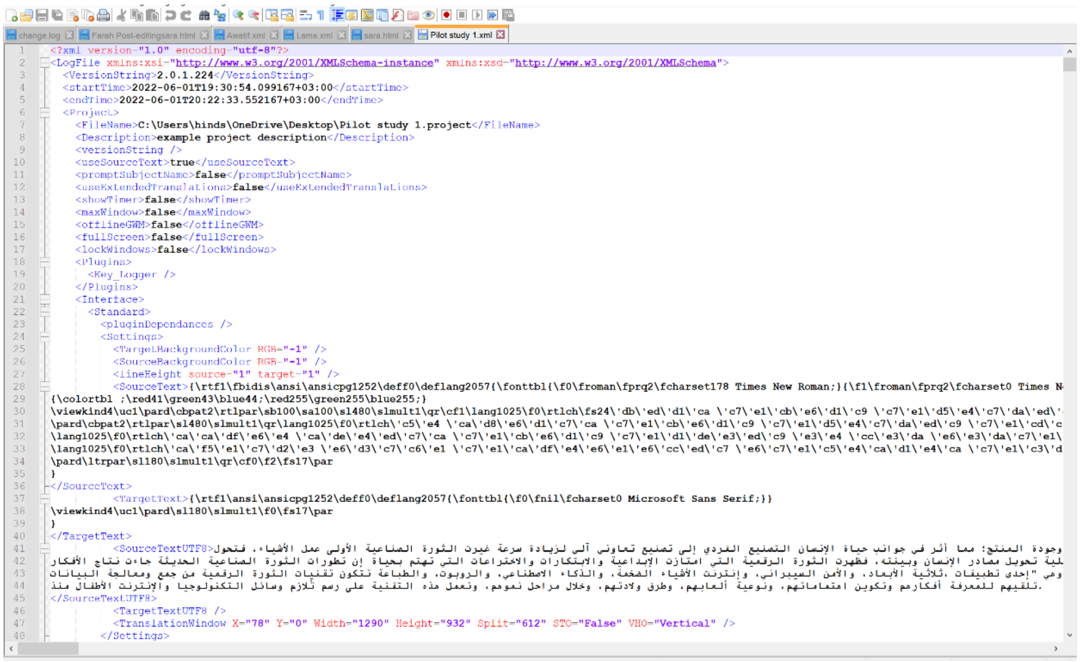

Figure 2. Translog-II Screenshot Logging Time

To address RQ2, which is concerned with investigating which of the two approaches requires more effort, the researchers collected data about the participants' keystrokes using Translog-II software, saving it as an .xml file. The "insert" and "delete" keystrokes of the participants in each approach were utilized as an indication of their effort during the translation process (Béchara et al., 2021; see Figure 3). In this study, time was utilized not only to measure speed but also as an effort indicator (Koponen et al., 2012).

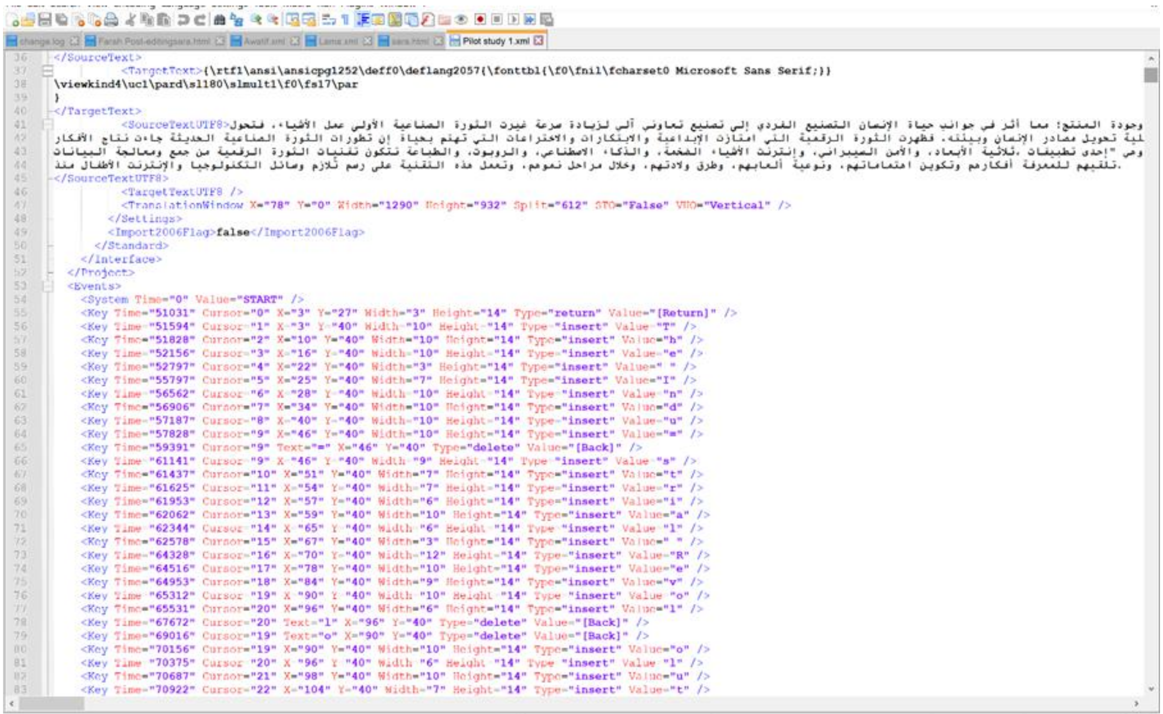

Figure 3. Translog-II Screenshot Recording "Insert" and "Delete" Keystrokes

IV. Results

A. Participants' Speed in TFS and PE

The researchers used time as an objective and easily quantifiable measure to answer RQ1. This is because time is considered a simple numerical measure rather than a complex function (Tatsumi, 2010). To obtain the necessary data, the researchers referred to the participants' start and end times recorded in Translog-II, as shown in Figure 3. Subsequently, the researchers calculated the time spent on each approach by subtracting the start time from the end time.

Figure 4 illustrates that there is a noticeable pattern among the participants in terms of the difference in speed between the two approaches. There was a significant difference between the PE time and the TFS time for most participants, with the former being considerably lower than the latter. For example, the lowest value for PE time was 5 minutes (P8), while the highest value was 63 minutes (P6). In contrast, the lowest value for TFS time was 53 minutes (P4), while the highest value was 103 minutes (P3).

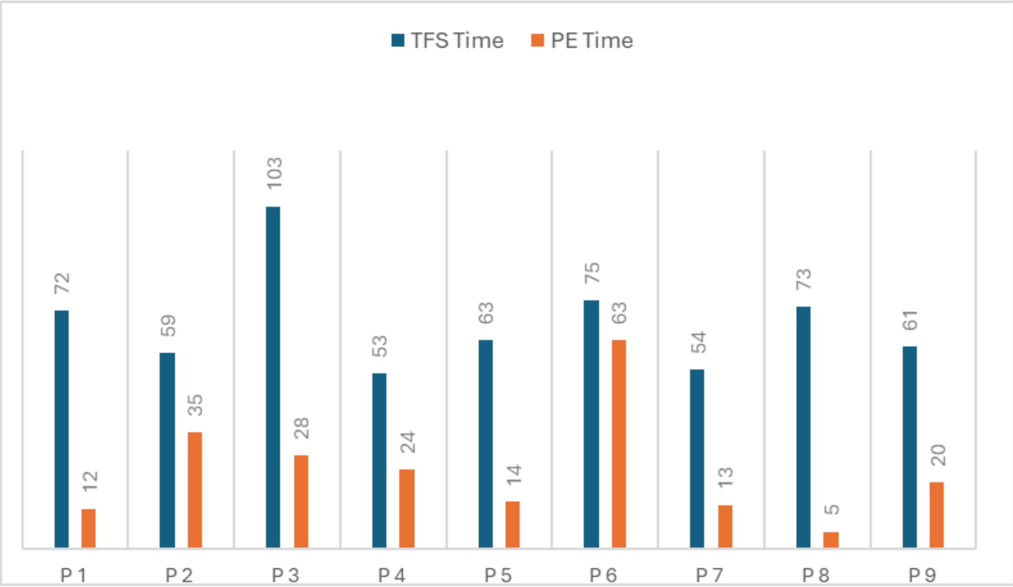

Figure 4. Participants' Times in TFS and PE

The speed difference between the two approaches also varied considerably among the participants. For instance, P8 took only 5 minutes in PE but required 73 minutes in TFS. Thus, their PE time was only 6.8% of their TFS time; this was the largest time difference among all participants. In contrast, P6 had the smallest time difference between the two approaches; they required 75 minutes in TFS and 63 minutes in PE, which was the longest PE time among all participants and was 84% of their TFS time.

In general, these results show that PE is substantially faster than TFS. However, the speed rate can vary among individuals. One possible explanation for the faster speed of PE is that, in PE, the translator's role is to examine and correct the existing translation instead of translating the entire text on their own, as in TFS. However, it is important to note that the quality of the final translation is also a critical factor in evaluating the effectiveness of PE.

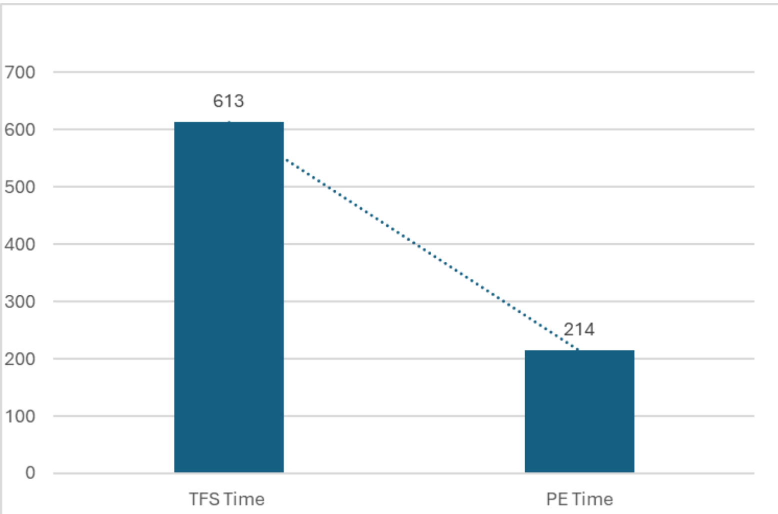

Figure 5. Total Time in TFS and PE

The total time spent on TFS was 613 minutes, whereas the time spent on PE was 214 minutes, indicating a considerable difference in the overall time taken between the two approaches. Specifically, the results show that the PE approach was 65.11% faster than the TFS approach. Furthermore, the *p*-value obtained was 0.00003, clearly indicating that the difference between the TFS and PE times was statistically significant at the conventional significance level of $p < 0.05$. This suggests that random chance is not likely to explain the observed difference in means, and it may have been caused by an actual difference between the two groups. These findings provide evidence to support the hypothesis that performing PE on CAT tool-generated text helps with expediting the process, which results in higher productivity rates.

B. Participants' Effort in TFS and PE

The analysis of the participants' effort in the TFS and PE approaches confirms the validity of the study's second hypothesis, which proposes that PE requires significantly less effort than traditional human translation. The researchers used Translog-II to collect data on the number of keystrokes logged (insertion and deletion) for each participant as well as the recorded time spent on each approach by seven of the participants. Figure 6 presents the number of keystrokes for each approach, facilitating a closer examination of the effort required for each one. The results show that the total number of keystrokes was significantly higher in TFS than in PE. Specifically, the number of insertions and deletions in TFS was more than 23,258 strokes and more than 4,133 strokes, respectively, accounting for approximately 15.1% of the total keystrokes in TFS. In contrast, the number of insertions and deletions in PE was more than 2,527 strokes and more than 1,236 strokes, respectively, accounting for approximately 32.8% of the total keystrokes in PE. Thus, the total number of PE keystrokes was only 13.74% of the total number of TFS keystrokes, indicating that PE required significantly less technical effort than TFS. Furthermore, PE required less time than TFS, indicating that PE also required less temporal effort.

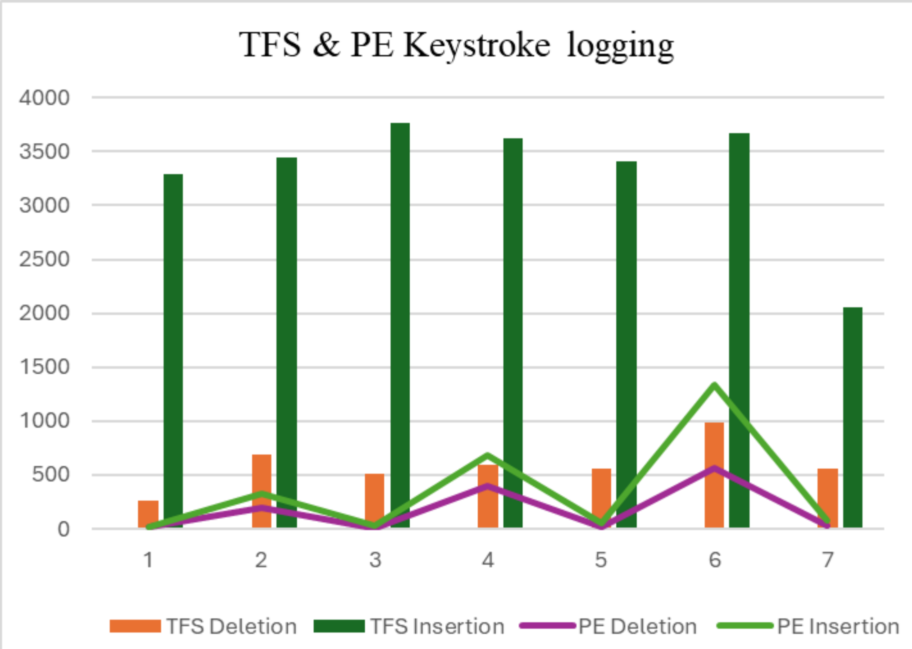

Figure 6. Participants' TFS and PE Keystroke Logging

Analyzing the time and keystrokes of each participant in the two approaches showed noteworthy discrepancies. For instance, a comparison of participants P1 and P2 in both approaches revealed significant differences. In the TFS approach, P1 recorded an approximate keystroke rate of 49.3 strokes per minute, while P2 recorded a very high rate of 69.98 strokes per minute. However, in the PE approach, P1 recorded a slower rate of 3.5 strokes per minute, while P2 recorded 14.9 strokes per minute. Thus, P2 displayed a 30% speed advantage over P1 in TFS and a substantial 76.5% advantage in PE. These patterns of differing speeds were consistently observed among all participants in the study. This finding indicates a notable performance difference between the participants in both approaches, and it may be worth exploring the factors contributing to these differences in performance.

C. Correlations Between Participants' Keystrokes and Time

The correlation between the variables analyzed in the previous section, specifically the participants' number of keystrokes and time taken in both approaches, was observed and identified to determine the level of effort required. A moderate positive correlation ($r = 0.40$) was found between the participants' keystrokes and time taken in TFS, as demonstrated in Figure 7. Additionally, a strong positive correlation ($r = 0.85$) was found between the participants' keystrokes and time taken in PE, as illustrated in Figure 8. This implies that, in PE, a higher number of keystrokes (indicating greater technical effort) tend to result in more time expended by the translator (indicating greater temporal effort).

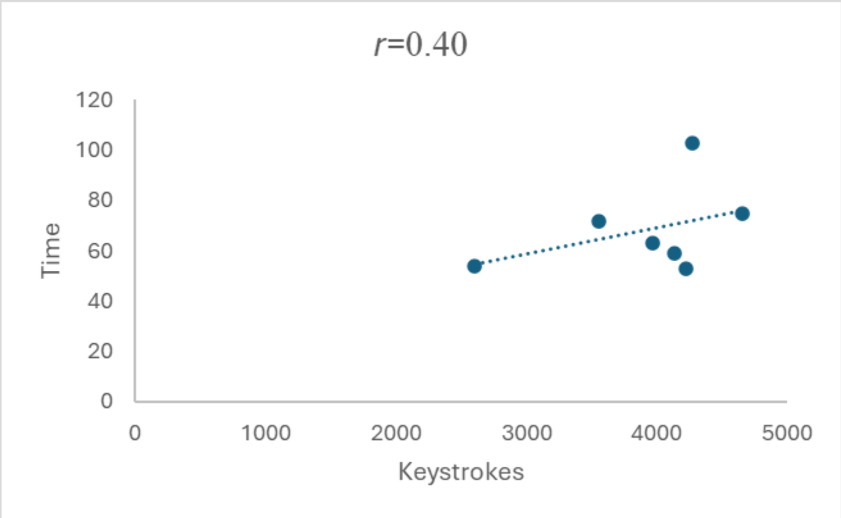

Figure 7. Correlation Between Participants' Keystrokes and Time Taken in TFS

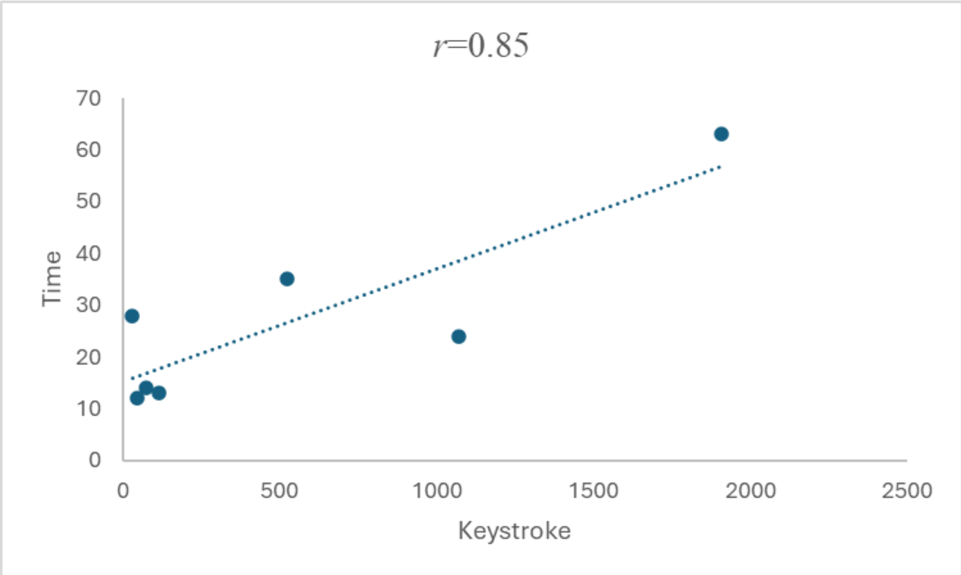

Figure 8. Correlation Between Participants' Keystrokes and Time Taken in PE

V. Discussion and Conclusion

This study revealed a distinct time difference between the two approaches, showing that PE was significantly faster than TFS by 65.1%. Specifically, PE required only 214 minutes, whereas TFS required 613 minutes. Notably, the range varied, from the quickest TFS completion time being 53 minutes to the longest being 103 minutes. In contrast, the lowest PE time was only 5 minutes (i.e., 9.4% of the lowest TFS time), while the highest PE time was 63 minutes (i.e., 61.1% of the highest TFS time). These results suggest that PE could significantly increase productivity and help the translation industry keep up with the substantial increase in translation volumes in a timely manner. These findings align with previous research. Studies by Vazquez et al. (2013), Läubli et al. (2013), and Aranberri et al. (2014) also identified differences in PE time and TFS time across different translators. Similarly, Haji Sismat (2016) and Samman (2022) noted varying efficiency levels between these methods; this is in line with the present study's outcomes. Building on this, the present study's findings highlight the significance of high-quality translations produced through TM or MT in achieving promising speed-related outcomes. Furthermore, the involvement of translators with PE proficiency played a significant role in contributing to these beneficial results. This finding supports the overall pattern arising from the literature, namely that PE can increase the productivity of translators in terms of speed. Koponen (2016a) obtained similar findings, noting that speed and productivity could be affected by conditions related to the high quality of TM or MT translation and post-editors' familiarity with the tools and processes involved in the PE task.

Regarding technical effort, the results indicate that translators required significantly more technical effort in TFS than in PE, with an 86.3% increase. Temporal effort was also found to be much lower in PE than in TFS, with the effort required in PE being only 35% of the effort required in TFS. These findings suggest that PE requires less effort both technically and temporally compared to TFS. Furthermore, it was observed that there were significant variations in the number of keystrokes per minute logged by different participants in each approach. For example, P1 logged approximately 49.3 keystrokes per minute in TFS, but only 3.5 keystrokes per minute in PE. Similarly, P2 logged 69.98 keystrokes per minute in TFS but only 14.9 keystrokes per minute in PE. These differences suggest that fewer keystroke logs and longer time could indicate a higher cognitive effort required to detect errors and plan the corresponding corrections. Overall, the results of this study indicate that PE requires less effort than TFS in both technical and temporal aspects. However, it is important to note that cognitive effort may vary among different participants and may require further investigation.

Based on this observation, the researchers found interesting correlations between the participants' keystrokes and time taken in TFS, which had a moderate positive correlation (r =0.40). Furthermore, the correlation between the participants' keystrokes and time taken in PE had a strong positive correlation (r =0.85). This indicates that, in PE, the higher the keystrokes (technical effort), the more time a translator takes (temporal effort). Koponen (2016b), whose study on PE and the effort involved yielded similar results, asked whether MT PE is worth the work. Other researchers found that editors vary considerably in terms of editing speed (Plitt & Masselot, 2010) and that they seem to differ more in terms of how much time they take and how many keystrokes they use rather than how many changes they make to the text. These studies also discovered that editors have different styles: Being fast or using several keystrokes does not necessarily mean more changes, and editors approach PE differently.

The increase in translation volumes and the imperative to enhance productivity have resulted in a growing need for translation technologies. This emphasizes the necessity for translators to acquaint themselves with CAT tools and understand how to utilize them effectively through PE (Bowker, 2015). Nevertheless, there is a pressing need for enhanced local training in CAT tools. A recent study (Al-Rumaih, 2021) highlighted deficiencies in CAT tool courses in Saudi Arabia, with students lacking the requisite technical expertise to navigate the demands of the professional sphere. However, Kiraly (2014) observed the following:

> Translator competence does not primarily refer to knowing the correct translation for words, sentences, or even texts. It does entail being able to use tools and information to create communicatively successful texts that are accepted as good translations within the community concerned. (p. 13)

This statement is particularly true at present, with advanced translation technologies being developed while the demand for them rises. Furthermore, limited experience with CAT tools will not allow translation students to gain a realistic understanding of the functioning of these tools (Bowker, 2015). This can also be applied to PE training, which is an intrinsic part of using translation technology in general; as mentioned before, these technologies cannot replace human translators and will always require human intervention to match the quality of human-edited text. Nevertheless, nearly all translators—especially local translators—need to receive training in PE. In a study on PE, Koby suggested that "the translator must be trained in post-editing" (Krings & Koby, 2001, p. 12). Furthermore, McElhaney and Vasconcellos (1988) argued that, since translation and PE are varying processes, translators are most suitable for undertaking this task, as they can identify linguistic errors and have rich knowledge about cross-language transfer. Therefore, in line with Al-Rumaih's (2021) implication, CAT tool course plans should be reviewed yearly to ensure they are up to date with technological advancements. Furthermore, PE training should be incorporated into CAT tool course. Workshops for students should sufficiently cover translation technology and ensure the practical use of these tools, in addition to providing PE training.

REFERENCES

[1] Allen, J. (2001). Postediting: An integrated part of a translation software program. *Language International Magazine*, *13*(2), 26–29.

[2] Allen, J. (2003). Post-editing. In H. Somers (Ed.), *Computers and translation: A translators guide* (p. 35). John Benjamins.

[3] Al-Rumaih, L. A. (2021). The integration of computer-aided translation tools in translator-training programs in Saudi universities: Toward a more visible state. *Arab World English Journal for Translation & Literary Studies*, *5*(1). http://dx.doi.org/10.2139/ssrn.3802984

[4] Aranberri, N., Labaka, G., de Ilarraza, A. D., & Sarasola, K. (2014). Comparison of post-editing productivity between professional translators and lay users. In *Proceedings of the 11th Conference of the Association for Machine Translation in the Americas* (pp. 20–33).

[5] Arenas, A. G. (2019). Pre-editing and post-editing. In *The Bloomsbury companion to language industry studies* (p. 333). Erik Angelone, Maureen Ehrensberger-Dow and Gary Massey (eds.) Bloomsbury Academic, 2020. https://doi.org/10.5040/9781350024960.0019

[6] Automatic Language Processing Advisory Committee. (1966). *Machines: Computers in translation and linguistics* (Publication 1416). Division of Behavioral Sciences, National Academy of Sciences, National Research Council.

[7] Balling, L. W., Carl, M., & O'Brian, S. (Eds.). (2014). *Post-editing of machine translation: Processes and applications*. Cambridge Scholars Publishing.

[8] Béchara, H., Orăsan, C., Parra Escartín, C., Zampieri, M., & Lowe, W. (2021, September). The role of machine translation quality estimation in the post-editing workflow. *Informatics*, *8*(3), 61. https://doi.org/10.3390/informatics8030061

[9] Boase-Beier, J. (2011). Stylistics and translation. In K. Malmkjaer & K. Windle, *The Oxford handbook of translation studies* (pp. 71–82). Oxford University Press.

[10] Bowker, L. (2002). *Computer-aided translation technology: A practical introduction*. University of Ottawa Press. https://doi.org/10.1017/CBO9781107415324.004

[11] Bowker, L. (2015). Computer-aided translation: Translator training. In S. Chan (Ed.), *Routledge encyclopedia of translation technology* (pp. 88–104). Routledge.
[12] Bowker, L. (2020). Terminology management. In *The Bloomsbury companion to language industry studies*. London: Bloomsbury. (pp. 261–283).
[13] Center for Research and Innovation in Translation and Translation Technology. (n.d.). *Translog II*. Retrieved September 22, 2022, from https://sites.google.com/site/centretranslationinnovation/translog-ii
[14] Chan, S. (Ed.). (2014). *Routledge encyclopedia of translation technology* (1st ed.). Routledge.
[15] Chesterman, A. (2004). Beyond the particular. In A. Mauranen & P. Kujamäki (Eds.), *Translation universal: Do they exist?* John Benjamins Publishing Company.
[16] Christensen, T. P., & Schjoldager, A. (2010). Translation-memory (TM) research: What do we know and how do we know it? *Hermes – Journal of Language and Communication in Business*, *23*(44), 89–101.
[17] Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, *20*(1), 37-46.
[18] Daems, J., Vandepitte, S., Hartsuiker, R., & Macken, L. (2017). Translation methods and experience: A comparative analysis of human translation and post-editing with students and professional translators. *Meta: Journal des traducteurs* [Meta: Translators' Journal], *62*(2), 245–270.
[19] De Sousa, S. C., Aziz, W., & Specia, L. (2011, September). Assessing the post-editing effort for automatic and semi-automatic translations of DVD subtitles. In *Proceedings of the International Conference Recent Advances in Natural Language Processing 2011* (pp. 97–103).
[20] Densmer, L. (2014). *Light and full MT post-editing explained*. Rws Moravia. Retrieved June 2, 2020, from http://info.moravia.com/blog/bid/353532/Light-and-Full-MT-PostEditing-Explained Web
[21] Do Carmo, F., & Moorkens, J. (2020). Differentiating editing, post-editing and revision. In *Translation revision and post-editing* (pp. 35–49). Routledge.
[22] Doyle, M. S. (2003). Translation pedagogy and assessment: Adopting ATA's framework for standard error marking. *The ATA Chronicle*, *32*(11), 21–28.
[23] Dragsted, B. (2004). *Segmentation in translation and translation memory systems: An empirical investigation of cognitive segmentation and effects of integrating a TM system into the translation process*. Samfundslitteratur.
[24] Elming, J., & Balling, L. W. (2014). Investigating user behaviour in post-editing and translation using the CASMACAT workbench. In *Post-editing of machine translation: Processes and applications.* Cambridge Scholars Publishing: Newcastle upon Tyn (p. 147).
[25] El-Zeini, N. T. (1994). *Criteria for the evaluation of translation: A pragma-stylistic approach*. Cairo University.
[26] Espan˜a-Bonet, C., & Costa-jussa, M. R. (2016). Hybrid machine translation overview. In M. Costa-jussa, R. Rapp, P. Lambert, K. Eberle, R. Banchs, & B. Babych (Eds.), *Hybrid approaches to machine translation* (pp. 1–24). Springer.
[27] Esselink, B. (2000). *A practical guide to localization*. John Benjamins.
[28] European Standards for Translation Services. (2004). *Translation services – Service requirements* (Draft prEN15038). Retrieved May 2011 from http://web.letras.up.pt/egalvao/prEN-15038.pdf
[29] Farahzad, F. (1992). Testing achievement in translation classes. In C. Dollerup & A. Loddegaard (Eds.), *Teaching translation and interpreting: Training, talent and experience* (pp. 271–278).
[30] Fiederer, R., & O'Brien, S. (2009). Quality and machine translation: A realistic objective. *The Journal of Specialised Translation*, *11*(11), 52–74.
[31] García, I. (2006). Translators on translation memories: A blessing or a curse? In *Translation technology and its teaching* (pp. 97–105).
[32] Garcia, I. (2009). Beyond translation memory: Computers and the professional translator. *The Journal of Specialised Translation*, *12*(12), 199–214.
[33] Garcia, I. (2010). Is machine translation ready yet? *Target – International Journal of Translation Studies*, *22*(1), 7–21.
[34] García, I. (2012). A brief history of postediting and of research on postediting. *Revista Anglo Saxonica*, *3*(3), 291–310.
[35] Gaspari, F., Almaghout, H., & Doherty, S. (2015). A survey of machine translation competences: Insights for translation technology educators and practitioners. *Perspectives*, *23*(3), 333–358.
[36] Glesne, C. & Peshkin, A. (1992). *Becoming qualitative researchers: An introduction*. Longman.
[37] Gray, D. E. (2004). *Doing research in the real world*. SAGE Publications.
[38] Guerberof, A. (2008). Productivity and quality in the post-editing of outputs from translation memories and machine translation. *Localisation Focus – The International Journal of Localisation*, *7*(1), 11–21.
[39] Guerberof, A. (2009). *Productivity and quality in MT post-editing* [conference paper]. Universitat Rovira I Virgilli, Spain. Retrieved July 17, 2014, from http://www.mt-archive.info/MTS-2009-Guerberof.pdf
[40] Haji Sismat, M. A. B. (2016). *Quality and productivity: A comparative analysis of human translation and post-editing with Malay learners of Arabic and English* [Doctoral dissertation, University of Leeds].
[41] Hargrave, J., & Savourel, Y. (1997). *Machine assisted translation tools*. Retrieved September 1, 2022, from https://patents.google.com/patent/US5724593A/en
[42] House, J. (1997). *Translation quality assessment: A model revisited*. Gunter Narr.
[43] Hu, K., & Cadwell, P. (2016). A comparative study of post-editing guidelines. *Baltic Journal of Modern Computing*, *4*(2), 346–353.
[44] Hutchins, J. (2000). *Yehoshua Bar-Hillel* (Vol. 97). John Benjamins Publishing.
[45] Hutchins, W. J. (1986). *Machine translation: past, present, future*. Ellis Horwood.
[46] Hutchins, W. J., & Somers, H. L. (1992). *An introduction to machine translation* (Vol. 362). Academic Press.
[47] Karwacka, W. (2013). Constructing translation research with the use of keystroke logging—A case study. In *Forum Filologiczne Ateneum* (p. 93).
[48] Kay, M. (1980). The proper place of men and machines in language translation. *Machine Translation*, *11*(1–2), 1–20. https://doi.org/10.1023/A:1007911416676

[49] Kiraly, D. (2014). *A social constructivist approach to translator education: Empowerment from theory to practice*. Routledge.
[50] Koponen, M. (2016a). *Machine translation post-editing and effort: Empirical studies on the post-editing process*. A PhD thesis, University of Helsinki: Helsinki.
[51] Koponen, M. (2016b). Is machine translation post-editing worth the effort? A survey of research into post-editing and effort. *The Journal of Specialised Translation*, *25*, 131–148.
[52] Koponen, M., Aziz, W., Ramos, L., Specia, L., Rautio, J., Gonzalez, M., … & Nyrkkö, S. (2012). Post-editing time as a measure of cognitive effort. In *AMTA 2012 Workshop on Post-editing Technology and Practice*.
[53] Koponen, M., & Mossop, B. (2021). *Translation revision and post-editing*. Routledge: London. https://doi.org/10.4324/9781003096962
[54] Krings, H. P., & Koby, G. S. (2001). *Repairing texts: Empirical investigations of machine translation post-editing processes*. Kent State University Press.
[55] Läubli, S., Fishel, M., Massey, G., Ehrensberger-Dow, M., Volk, M., O'Brien, S., & Specia, L. (2013). *Assessing post-editing efficiency in a realistic translation environment*. In Proceedings of the 2nd Workshop on Post-editing Technology and Practice, Nice, France.
[56] Lauffer S. (2008). The translation process: An analysis of observational methodology. *Cadernos de Tradução*, *9*(1), 57–75. Retrieved September 1, 2022, from: http://journal.ufsc.br/.
[57] Martínez-Mateo, R., Montero Martínez, S., & Moya Guijarro, A. J. (2017). The modular assessment pack: A new approach to translation quality assessment at the Directorate General for Translation. *Perspectives*, *25*(1), 18–48.
[58] McElhaney, T., & Vasconcellos, M. (1988). The translator and the postediting experience. *Technology as Translation Strategy*, *2*, 140–148. https://doi.org/10.1075/ata.ii.28mce
[59] Melby, A. (1979). ITS : Interactive translation system. Interactive Translation System. *International Conference on Computational Linguistics*, *5*(1), 234–241. Retrieved September 1, 2022, from file:///Users/SK/Documents/MyResearch/Library/articles/C80-1064.pdf
[60] Melby, A. (1981). Translators and machines – Can they cooperate? *Meta: Journal Des Traducteurs* [Meta: Translators' Journal], *26*(1), 23. https://doi.org/10.7202/003619ar
[61] Melby, A. (1982). Multi-level translation aids in a distributed system. In *Proceedings of the Ninth International Conference on Computational Linguistics* (pp. 215–220). Retrieved September 1, 2022, from http://ww2.cs.mu.oz.au/acl/C/C82/C82-1034.pdf
[62] Melby, A. (1992). The translator workstation. In John Newton (Ed.), *Computers in translation: A practical appraisal* (pp. 147–165). https://doi.org/10.1075/ata.vii.14me
[63] Melby, A. K. (1983, February). Computer-assisted translation systems: The standard design and a multi-level design. In *First Conference on Applied Natural Language Processing*, pages 174–177, Santa Monica, California, USA. Association for Computational Linguistics.
[64] Moorkens, J., Toral, A., Castilho, S., & Way, A. (2018). Translators' perceptions of literary post-editing using statistical and neural machine translation. *Translation Spaces*, *7*(2), 240–262.
[65] Munday, J. (2016). *Introducing translation studies: Theories and applications*. Routledge.
[66] Newmark, P. (1995). *A textbook of translation*. Longman.
[67] Nida, E. A. (1964). *Toward a science of translating: With special reference to principles and procedures involved in Bible translating*. Brill Archive.
[68] Nida, E. A. (2001, November). *Language and culture-contexts in translating*. Shanghai Foreign Language Education Press.
[69] Nord, C. (2001). *Translating as a purposeful activity-functionalist approaches explained*. Shanghai Foreign Language Education Press.
[70] O'Brien, S. (2007). An empirical investigation of temporal and technical post-editing effort. *Translation and Interpreting Studies – The Journal of the American Translation and Interpreting Studies Association*, *2*(1), 83–136.
[71] O'Brien, S. (2010). Introduction to post-editing: Who, what, how and where to next? In *Proceedings of the 9th Conference of the Association for Machine Translation in the Americas: Tutorials*.
[72] O'Brien, S. (2011). Towards predicting post-editing productivity. *Machine Translation*, *25*(3), 197–215.
[73] Phelan, M. (2017). Analytical assessment of legal translation: A case study using the American Translators Association framework. *The Journal of Specialised Translation*, *27*, 89–210.
[74] Plitt, M., & Masselot, F. (2010). A productivity test of statistical machine translation post-editing in a typical localisation context. *The Prague Bulletin of Mathematical Linguistics*, *93*, 7–16.
[75] Samman, H. M. (2022). *Evaluating machine translation post-editing training in undergraduate translation programs—An exploratory study in Saudi Arabia* [Doctoral dissertation, University of Southampton].
[76] Schäfer, F. (2003). MT post-editing: How to shed light on the "unknown task". Experiences at SAP. In *EAMT Workshop: Improving MT through other language technology tools: Resources and tools for building MT* (p. 3).
[77] Secară, A. (2005). Translation evaluation: A state of the art survey. In *Proceedings of the eCoLoRe/MeLLANGE Workshop* (pp. 39–44). St. Jerome Publishing.
[78] Shakir, R. M. F., & Al-Ali, K. K. (2021). Post-editing as a creative tool in improving the quality of the product of translation students. *Adab Al-Basrah*, *97*(1), 1227-1235.
[79] Sharma, V. (2015). The relevance of addition, omission and deletion (AOD) in translation. *International Journal of Translation*, *27*(1–2), 1–12.
[80] Somers, H. (Ed.). (2003). *Computers and translation: A translator's guide*. John Benjamins.
[81] Tatsumi, M. (2010). *Post-editing machine translated text in a commercial setting: Observation and statistical analysis* [Doctoral dissertation, Dublin City University].
[82] Temnikova, I. (2010, May). Cognitive evaluation approach for a controlled language post-editing experiment. In *Proceedings of the Seventh International Conference on Language Resources and Evaluation* (LREC'10).
[83] Terninko, J. (2018, January). *2018 European language industry survey*. European Language Industry Association. Retrieved September 1, 2022, from https://elia-association.org/2018/01/2018-european-language-industry-survey/

[84] Tetnowski, J. (2015). Qualitative case study research design. *Perspectives on Fluency and Fluency Disorders*, *25*(1), 39–45.
[85] Vazquez, L. M., Vazquez, S. R., & Bouillon, P. (2013). Comparing forum data post-editing performance using translation memory and machine translation output: A pilot study. In *Proceedings of Machine Translation Summit XIV* (pp. 249–256).

Hind S. Alsaif earned a Master of Translation from Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia in 2022.

She has extensive professional experience as a translator, with a particular focus on translation technologies, Computer-Assisted Translation (CAT), and specialized translation fields. Prior to her current position, she held various roles that expanded her expertise in both the practical and academic aspects of translation. She currently works at the Decision Support Centre-Royal Court, located in Riyadh, Saudi Arabia.

Alsaif is actively involved in professional societies related to her field. Her contributions to translation studies and related technologies underscore her commitment to advancing this field. Her work has been recognized in several professional circles, although specific awards and committee roles are not mentioned.

Ebtisam S. Aluthman earned a Ph.D. in Applied Linguistics from the University of Manchester, United Kingdom. Additionally, she obtained her Master's in Applied Linguistics and a Bachelor of Arts in English Language from Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

She is currently working as the Secretary-General of the Princess Nourah bint Abdulrahman Award for Women's Excellence and Advisor to the President of Princess Nourah bint Abdulrahman University. Dr. Aluthman is an esteemed member of several academic societies and is evidently interested in the intersection of computing and linguistics. This may involve developing or studying software and tools that support language learning and processing, such as natural language processing (NLP) applications.