
HERMES

**Journal of Language and
Communication in Business**

No. 66 – 2026

Published by the School of Communication and Culture, Aarhus University

ISSN: 1903-1785

Rouhullah Nemati Parsa

How Training Enhances Post-Editing Speed in Iranian BA Translation Students: A Case of Google-Translated Financial Texts

Abstract

The increasing adoption of machine translation post-editing (PE) in the language industry underscores the need for effective training to improve PE speed. However, empirical evidence on the effectiveness of PE training remains limited, particularly for the English–Persian language pair. This study investigated the effect of a guideline-based training program on PE speed among undergraduate English Translation students in Iran. Participants were randomly assigned to either an experimental group ($n = 10$), which received a structured 16-session PE training program based on the TAUS (2017) PE guidelines, or a control group ($n = 10$), which received no formal PE instruction. Before the intervention, participants completed a 15-item Likert-scale questionnaire assessing their perceptions of PE speed and efficiency. Following the training period, participants post-edited the same 400-word English–Persian financial text, pre-translated using Google Translate, on the MateCat platform. PE speed was measured using actual task completion time and words per minute (WPM), calculated from the MateCat editing log. The experimental group completed the task significantly faster than the control group ($M = 44.20$ vs. 68.90 min), Welch's $t(15.61) = -7.28$, $p < .001$, $d = 3.25$, and achieved a significantly higher WPM rate (9.20 vs. 5.89), Welch's $t(14.97) = 7.13$, $p < .001$, $d = 3.19$. Participants in the experimental group also reported significantly more positive perceptions of PE speed and efficiency, controlling for pre-test scores (ANCOVA, $p = .002$). These findings support the integration of guideline-based PE training into translator education.

Keywords

Financial Text; Google Translate; Machine Translation; Post-Editing; Translation Training

1 Introduction

Machine translation (MT) has gained widespread adoption across diverse fields due to its economic advantages. According to the Translation Automation User Society (TAUS) guidelines, “Machine translation (MT) with post-editing (PE) is rapidly becoming a standard industry practice” (TAUS, 2017, p. 5). This requires organizations to efficiently identify, qualify, train, and evaluate post-editors' performance. However, gaps remain between industry demands and translator education. In this context, González Pastor (2024) highlights the limited integration of translation technologies into Spanish translation and interpreting curricula. This gap is similarly evident in Iran, where both MT and PE training fall short of equipping students with the skills demanded by the evolving job market (Jovein et al., 2024).

Given the significant pedagogical gap between academic training and industry expectations, and the growing reliance on MT systems in professional translation workflows (Guerberof Arenas & Moorkens, 2019; Kenny, 2022; O'Brien, 2021), there is an urgent need to empirically investigate whether formal PE instruction can meaningfully enhance student performance. Speed, in particular, is a critical metric in professional PE contexts, as it directly affects productivity, cost-efficiency, and the marketability of translation graduates (De Almeida & O'Brien, 2010; Läubli et al., 2019). Although previous studies have reported that structured training can improve PE productivity and efficiency (Blagodarna, 2018; Samman, 2022; Toledo Báez, 2018), empirical research has examined the effects of such training interventions on PE speed only to a limited extent, particularly in underexplored language pairs such as English–Persian and in educational contexts, such as Iranian universities (Arjani & Jamshidiha, 2024; Asadi et al., 2023). As highlighted by Vashee (2013), MT is advantageous when expediency and timeliness take precedence over perfect translation quality, making it

* Rouhullah Nemati Parsa
Independent researcher
Rouhullah.parsa@gmail.com

well-suited for handling dynamic content that requires frequent updates. However, achieving Fully Automatic High-Quality Machine Translation (FAHQMT) remains a formidable challenge owing to the inherent intricacy of natural language and unsettled issues within MT technology, which confine its effectiveness to highly specialized domains (Koehn, 2009). This highlights the importance of human involvement in translation. In response to these limitations, PE has become a vital solution. While MT can produce acceptable translations within specific domains, it often struggles to capture the nuances of natural language. Human post-editing ensures accuracy, rectifies domain-specific terminology, and bridges the gap between machine-generated text and high-quality human translation. Research suggests that PE can, at times, be more efficient than starting from scratch, particularly for large volumes of content (Garcia, 2011). This is because MT excels at handling repetitive elements and basic structure, while human intervention through PE ensures accuracy and nuance (Koehn, 2009).

PE emerged in the 1950s as a response to the shortcomings of early MT systems (Bar-Hillel, 1960). Early definitions of PE, as proposed by Yngve (1954), emphasized the need for human intervention to improve the readability or comprehensibility of machine outputs. Over time, the definition of PE has evolved in tandem with advancements in MT technology. Allen (2001) defines it as the editing and correction of pre-translated texts generated by MT systems. Schäfer (2003) focuses on the goal of "polishing" raw MT output into a user-friendly format. These definitions underscore the collaborative nature of PE, where human expertise refines the initial MT draft. The burgeoning demand for translations, coupled with the limitations of traditional human translation methods, has fueled the rise of PE. As O'Brien (2011) notes, translators leverage PE to meticulously review and modify initial machine translations, ensuring they meet pre-established quality standards. This human-machine workflow streamlines the translation process by reducing the workload compared to purely human translation (Toral et al., 2018). PE has become an indispensable tool for addressing the ever-increasing need for translations, which traditional methods struggle to keep pace with (TAUS, 2017). The expanding adoption of MT necessitates improved human-machine collaboration and research on optimizing PE processes (Çetiner & İşısağ, 2019). While untrained individuals previously lacked the necessary skills for PE (Aranberri, 2017), advanced MT systems now produce initial drafts across various languages and domains, rendering PE an essential practice for professional translators.

2 Literature Review

This literature review is organized into three thematic blocks that collectively establish the foundation for the present study. The first section examines PE research and the role of training in developing PE competencies. The second section explores the specific challenges posed by financial translation and its intersection with PE. The third section analyzes PE research across language pairs, highlighting the under-researched status of certain language combinations. Together, these blocks converge to identify a critical gap at their intersection, which the present study aims to address.

2.1 Post-Editing and Training: Foundations, Efficiency, and Pedagogical Implications

PE outputs have become a significant area of investigation within Translation Studies, examined from diverse perspectives. Foundational work by Krings (2001) pioneered the empirical study of PE by analyzing the process of translating technical texts from English, French, and German using MT systems. Employing the think-aloud protocol with 52 participants, Krings acknowledged the method's intrusiveness while demonstrating its value in revealing the cognitive processes involved in PE.

Early explorations continued with Torrejón and Rico (2002), who investigated controlled-language scenarios by combining error analysis and pre-editing techniques to enhance MT quality. Temnikova (2010) further examined edit types, finding that PE output derived from simplified texts required less PE time than output from more complex source texts. Although controlled language operates on the source text rather than the PE task itself, both lines of research highlight the same underlying factor influencing PE effort: the complexity and predictability of the input. Together,

these studies provided an early empirical basis for later PE training research by underscoring the importance of understanding how source-text characteristics and MT error patterns affect editing effort, thereby supporting the inclusion of error awareness and editing strategies in structured PE training programmes.

A major thread of PE research compares its efficiency and cognitive demands with those of traditional human translation. Carl et al. (2011) demonstrated that PE could reduce translation time compared with manual translation, suggesting potential productivity gains. Koglin (2015), using eye-tracking and keystroke logging, concluded that PE generally required less cognitive effort than manual translation, particularly when editing machine-translated metaphors. Similarly, Jia et al. (2019) found that PE of neural machine translation (NMT) output imposed a lower cognitive load than translation from scratch, particularly for specialised texts. Jia and Sun (2023) further reported that PE was substantially less demanding than human translation when MT quality was high. However, the magnitude of this advantage depended on text complexity and MT performance.

The transition from Statistical Machine Translation (SMT) to NMT has also substantially influenced PE research. Koponen et al. (2019) reported that PE effort was significantly lower for NMT than SMT output. Likewise, Yamada (2019) identified notable differences in editing patterns between SMT and NMT systems. Highlighting professional applications, Läubli et al. (2019) demonstrated that NMT, when combined with translation memories and terminology databases, considerably accelerated financial translation even with limited domain adaptation.

Beyond productivity, researchers have increasingly examined translators' attitudes and learning experiences. Moorkens and O'Brien (2015) found that professional translators completed PE tasks more rapidly and held more positive attitudes than undergraduate students. Guerberof Arenas and Moorkens (2019) similarly reported that prior PE experience positively influenced translators' perceptions of MT. Harto et al. (2022) observed that structured PE instruction improved students' editing competence and attitudes toward PE.

Pérez (2024) argues that existing PE guidelines are either overly generic or excessively context-specific, limiting their pedagogical usefulness. To overcome these limitations, she proposes an integrated PE ecosystem that systematizes recurrent editing challenges and provides more consistent guidance for translators. Pérez's framework further highlights the need for empirical studies evaluating whether structured training can translate theoretical guidelines into measurable improvements in PE performance.

Sánchez Ramos (2025) investigates trainee translators' emotional responses to MT and PE using an emotional narrative methodology in which students wrote either "love letters" or "break-up letters" to MT. Overall, participants expressed more positive than negative attitudes, primarily associated with perceived efficiency and satisfaction with MT output. These findings are consistent with the earlier work of Koskinen and Ruokonen (2017) ("Love letters or hate mail?"), which likewise demonstrated that translators' emotional engagement with MT influences their willingness to adopt PE technologies in future professional practice. These studies therefore suggest that successful PE training should address not only technical competence but also learners' attitudes toward MT.

While much research, including Green et al. (2013) and Plitt and Masselot (2010), supports the productivity advantages of PE, more recent evidence advocates a more cautious interpretation. Terribile (2023) challenged the widespread assumption that PE is invariably faster than human translation by demonstrating substantial variability in translation speed and weak correlations between edit distance and productivity. These findings suggest that productivity gains cannot be attributed solely to the use of MT; rather, they are shaped by multiple interacting factors, including translator expertise, text type, MT quality, and the effectiveness of training. Consequently, examining structured training interventions is essential for understanding the conditions under which PE enhances translation efficiency. This perspective provides a natural transition to the domain-specific challenges discussed in the following section.

2.2 Financial Translation and Post-Editing: Domain-Specific Challenges and Opportunities

Financial texts pose unique challenges for both MT systems and human post-editors because of their specialised terminology, syntactic complexity, and stringent accuracy requirements. Google Translate, despite considerable advances in NMT, frequently produces financial translations requiring substantial human intervention (Ali & Alhebshi, 2026; Kriston, 2017). Research has shown that PE can nevertheless generate considerable productivity gains even in less-resourced domains when translators receive appropriate domain-specific training (Saricaoglu & Dinçer, 2025).

Multiple studies consistently confirm the positive contribution of PE training to student performance, demonstrating improvements in translation quality (Blagodarna, 2018, 2020; Gu et al., 2025), PE competence (Al Sammarraie et al., 2024; Blagodarna, 2020), error detection and correction (Emara, 2023; Zhang & Torres-Hostench, 2022), productivity (Blagodarna, 2018; Samman, 2022), processing speed and efficiency (Samman, 2022; Stasimioti & Sosoni, 2019), evaluative judgement and revision justification (Gu et al., 2025), critical awareness of MT (Zhang & Torres-Hostench, 2022), learner confidence (Blagodarna, 2018; Stasimioti & Sosoni, 2019), positive attitudes toward PE and AI-assisted translation (Al Sammarraie et al., 2024; Samman, 2022; Stasimioti & Sosoni, 2019), collaborative learning and self-reflection (Wang & Wang, 2021), emotional engagement with PE (Sánchez Ramos, 2025), and reading comprehension and classroom participation in EFL contexts (Kim & Cha, 2023). Importantly, these studies should not be interpreted as contradicting the previously identified research gap. Rather, existing research has primarily demonstrated improvements in translation quality, error awareness, and attitudes toward PE. In contrast, considerably fewer studies have isolated PE speed as the principal outcome variable, particularly in undergraduate educational settings and low-resource language pairs. Trained post-editors consistently demonstrate greater awareness of MT errors and produce higher-quality translations than untrained peers (Dede & Antonova-Ünlü, 2022). Training plays a crucial role in enhancing post-editors' awareness of MT errors and their ability to produce higher-quality translations compared with untrained peers (Dede & Antonova-Ünlü, 2022), particularly by developing the professional judgement needed to distinguish between segments requiring substantial revision and those that can be adequately addressed through minimal corrections (Al Sammarraie & Ghassemiazghandi, 2024).

2.3 Post-Editing Across Language Pairs

A critical examination of the PE literature reveals a pronounced concentration on high-resource language pairs. Seminal empirical investigations have predominantly focused on English-German (Koponen et al., 2019; Krings, 2001), English-French (Läubli et al., 2019), English-Spanish (Moorkens & O'Brien, 2015), English-Chinese (Jia et al., 2019), and English-Japanese (Yamada, 2019). This concentration largely reflects the commercial importance of these language combinations and the availability of extensive parallel corpora for MT development. Consequently, substantially less attention has been devoted to lower-resource language pairs, limiting the generalisability of current findings across diverse linguistic settings. Emerging research has begun to extend the study of PE to other language pairs, including Arabic (Al Sammarraie, 2025; Eljazouli & Azmi, 2024), Persian (Bastan et al., 2017; Pilevar & Faili, 2010), and low-resource or morphologically complex languages such as Kazakh and Hindi-Marathi (Deoghare et al., 2024; Rakhimova et al., 2024), as well as Manipuri and Mizo (Sanayai Meetei et al., 2020). However, these studies remain considerably fewer in number and narrower in scope than research on high-resource pairs, and English-Persian in particular remains substantially under-investigated.

Taken together, the reviewed literature demonstrates that PE has become an indispensable component of contemporary translation practice and that formal training can enhance translators' quality awareness, error detection skills, and overall performance. Nevertheless, four important research gaps remain. First, empirical evidence continues to be dominated by high-resource language pairs, leaving English-Persian substantially under-investigated. Second, while numerous studies have examined translation quality, attitudes, or cognitive effort, relatively few have isolated PE speed as the primary outcome variable. Third, despite the importance of financial translation, comparatively little

research has examined structured PE training for financial texts within academic settings. Finally, undergraduate translation students, who constitute the future professional workforce, remain considerably less represented than professional translators in empirical PE research.

The present study addresses these interconnected gaps by examining whether a PE training program significantly improves the PE speed of Iranian BA Translation students working with Google Translate English–Persian financial texts. By integrating the dimensions of language pair, specialised domain, pedagogical intervention, and productivity measurement within a single experimental design, the study extends current PE research while providing evidence with direct implications for translator education and curriculum development. Accordingly, the study addressed the following research question:

RQ: Does Post-editing training have a positive effect on increasing students' post-editing speed?

Post-editing speed:

PE speed cannot be considered a clear-cut measure, as it is influenced by the complex interplay of multiple factors, including translators' skills and experience, perceptions, language pairs, the quality of MT output, and domain-specific characteristics, among others. Although it is generally conceptualized as the rate of PE per hour, some scholars define it as the speed variance among participants, which can be quantified. For example, in the financial text type, Läubli et al. (2019) report average NMT post-editing speeds of 934 WPH (word per hour) for German–French and 495 WPH for German–Italian. Similarly, Toral et al. (2018) report an average PE speed of 685 WPH for NMT of literary texts. In the present study, PE speed was operationalized through two complementary measures: (1) an objective measure, based on task-completion time (minutes) and words per minute (WPM), recorded via the MateCat editing log, which served as the primary indicator used to answer the research question; and (2) a subjective measure, reflecting participants' self-rated perception of their own PE speed, obtained from a 15-item Likert-scale questionnaire administered before and after the PE task. The perceptual measure was included as a secondary, complementary index to capture whether training also altered participants' subjective sense of efficiency, independently of their actual timed performance. In line with TAUS (2017), our study was based on training post-editors and trainees in advance to reduce the required PE level, as proposed by the guidelines (p. 15).

PE speed can be measured in terms of quality improvement by assessing how efficiently human post-editors enhance the output of MT systems to meet a desired quality standard (Devatine & Abraham, 2024; O'Brien, 2011). This relationship is multifaceted, involving both temporal aspects (speed) and qualitative outcomes (improvement). Essentially, improvement refers to enhancing the quality of the text, which can be quantified using metrics such as fluency, accuracy, and adherence to the intended meaning (Costa et al., 2021; Ortiz-Boix & Matamala, 2017).

Improvement in PE primarily refers to quality enhancement, whereas speed reflects the efficiency with which that improvement is achieved. In this study, quality improvement is operationalized through reductions in PE effort and verified using objective metrics, such as the Translation Edit Rate (TER), and human quality assessment (Daems et al., 2017). TER measures the amount of editing that a human would have to perform to change a system output so it exactly matches a reference translation (Snover et al., 2006, p. 1). Speed is measured using productivity indicators (e.g., words per hour) and is interpreted only when quality or fit-for-purpose (Bowker 2019) thresholds are met. Under these conditions, higher speed indicates more efficient attainment of equivalent quality improvement, not improvement in quality itself (Alvarez-Vidal & Oliver, 2023).

3 Method

3.1 Research design

This study employed a randomized, between-subjects experimental design with a pre-test/post-test structure to evaluate the impact of a PE training intervention. Participants were randomly assigned to one of two independent groups: an experimental group, which received the sixteen-session PE training program, and a control group, which received no formal PE instruction. Following

completion of the intervention, both groups completed the same PE task under identical conditions, allowing between-group comparisons of the objective completion-time and WPM measures.

In addition, a 15-item Likert-scale questionnaire was administered to both groups before the training intervention (pre-test) and again after completion of the intervention (post-test) to assess participants' self-rated perceptions of PE speed and efficiency. This questionnaire constituted a secondary outcome measure and was conceptually distinct from the objective performance measures.

To preserve this distinction, all analyses of objective PE performance were conducted using the completion-time and WPM measures recorded in MateCat, whereas analyses of the questionnaire data were interpreted solely as evidence regarding participants' subjective perceptions of their own PE performance.

3.2 Pilot Study

Before the main investigation, a pilot study was conducted with five Iranian BA English translation students, drawn from the target population but excluded from the main sample, to refine the research instruments and assess procedural feasibility. The primary objectives were to evaluate the clarity, comprehensibility, and face validity of a fifteen-item Likert scale questionnaire assessing perceived PE speed and to test the practicality of the overall research protocol, including task instructions, text suitability, and time allocation.

Google Forms was selected as the data-collection platform because it is freely accessible, familiar to the student population, and allows for fully online administration without software installation. Additionally, it provides automatic time-stamping of responses, minimizes data-entry errors through direct digital input, and facilitates secure, centralized storage and easy export of data for statistical analysis.

Pilot participants completed the questionnaire, performed the PE task on the designated 400-word Google Translate financial text, and provided qualitative feedback. This feedback indicated that the questionnaire items were generally clear and relevant, leading to minor wording adjustments in two items to enhance precision. The overall research procedure, including task administration and timing (observed range approx. 45-75 minutes), was confirmed as feasible and understandable. Consequently, the pilot phase validated the suitability of the core methodology and instruments, allowing the main study to proceed with increased confidence in the procedural design.

3.3 Participants

Ethical approval for this study was obtained from the Research Ethics Board of the University of Zanjan, and written informed consent was secured from all participants after a detailed explanation of the research objectives and procedures. Participants were selected through convenience sampling from a pool of Iranian BA translation students enrolled in the mandatory Translation and Technology course during the Fall 2024 semester at the University of Zanjan. From the volunteers who met the inclusion criteria (BA translation major, enrollment in the specific course, English as a foreign language status, and completion of introductory translation practice courses), a final sample of 20 students ($N = 20$) was selected. This sample consisted of native Persian speakers aged 20–24 years, primarily fourth-year undergraduates, presumed to have limited prior formal training in PE, as indicated by the curriculum. The 20 participants were then randomly assigned to either an experimental group ($n = 10$; 6 males, 4 females) or a control group ($n = 10$; 5 males, 5 females) using a computer-generated random sequence. This random assignment procedure aimed to enhance internal validity by maximizing the likelihood of achieving baseline equivalence between the groups on potential confounding variables, rather than ensuring representativeness of a broader population beyond the specific course context.

Both groups were drawn from the same cohort and enrolled in the same mandatory Translation and Technology course and were therefore matched on curriculum exposure, instructor, and academic level. The sole planned difference between conditions was that the experimental group additionally received a 16-session PE training module embedded within the course, whereas the control

group completed the same course requirements without this dedicated PE component and received no other formal PE instruction (see Section 3.5).

3.4 Instruments

Two primary instruments were used for data collection in this study: the PE task materials/environment and a perception questionnaire.

3.4.1 Post-Editing Task Materials and Environment

The core instrument for assessing objective PE speed consisted of the following components:

- A. Source Text: A 400-word English financial news excerpt adapted from a contemporary article published in The Wall Street Journal (Goldfarb, 2024). The text was pre-translated using Google Translate and served as the source text for the PE task.
- B. Machine Translation Output: The corresponding Persian translation of the source text, generated using the publicly available Neural Machine Translation engine, Google Translate.
- C. MateCat Platform

Participants performed the PE task in the specific Computer-Assisted Translation (CAT) tool environment of MateCat. (<https://www.matecat.com>). Matecat was selected as the CAT tool for this experiment, primarily for its integrated editing log functionality and user-friendly interface. As a web-based platform, Matecat facilitated participant access without requiring prior software installation and enabled centralized monitoring of the PE process. The editing log captures crucial metrics, including: (1) edit time per segment; (2) total job completion time; (3) average edit time across the task; and (4) an indicator derived from fuzzy match algorithms and the TER metric reflecting the proportion of edited words (Snover et al., 2006). The English source text was loaded into the tool's source segments, and Google Translate was used. Persian output was pre-populated into the target segments. Participants were instructed to use the CAT tool's standard editing features to modify the target segments, aiming to achieve the quality specified by the TAUS 2017 guidelines. Participants had access to standard features within the tool (e.g., spell check and concordance search, where applicable). They were permitted to use external resources, such as online dictionaries or terminology tools. This flexibility was consistent with realistic PE scenarios. The primary data collected for this task were the total times each participant took to finalize all segments.

3.4.2 Perception Questionnaire

The secondary instrument was a researcher-constructed questionnaire designed to assess participants' perceptions of PE speed, particularly regarding the perceived effects of the training intervention. The questionnaire consisted of two sections: (1) demographic information (age, gender), and (2) a 15-item questionnaire assessing perceptions related to PE speed factors. Participants answered a 15-item, 5-point Likert-scale questionnaire at both pre-test and post-test (1 = not at all; 5 = completely). The questionnaire was administered electronically via Google Forms following the completion of the PE task.

3.5 Research Procedure

The study was conducted during a Translation and Technology course offered from September 23, 2024, to January 19, 2025. Participants were randomly assigned to either an experimental group ($n = 10$) or a control group ($n = 10$). The experimental group received a structured 16-session PE training program integrated into the Translation and Technology course and based on the TAUS (2017) post-editing guidelines. In contrast, the control group did not receive formal PE training and continued with the regular Translation and Technology course throughout the study period. The PE training aimed to improve students' PE speed while developing their post-editing competence and productivity. The sessions included instruction on the principles of PE, the identification and correction of MT errors, the application of the TAUS (2017) Post-Editing Guidelines, and guided post-editing practice

using the MateCat computer-assisted translation platform. Participants also completed error analysis activities, monitored their productivity through time-effort analysis and words-per-minute (WPM) calculations, and received feedback on their performance. The training concluded with a 1,000-word post-editing project in which students applied the skills they had acquired while monitoring their PE speed.

Before the training program began, all participants completed a 15-item Likert-scale questionnaire to assess their perceptions of post-editing speed and efficiency (pre-test). Following the training period, all participants attended a one-hour controlled experimental session in which they post-edited the same 400-word English–Persian financial text, pre-translated using Google Translate, on the MateCat platform. A maximum time limit of 90 minutes was established solely for scheduling purposes. As no participant approached this limit (completion times ranged from 35 to 85 minutes), it served only as a practical safeguard and was not treated as a performance benchmark in the analyses. During the post-editing task, participants were permitted to consult specialised dictionaries and other online resources to simulate authentic professional post-editing practice. PE speed was measured using participants' actual task completion time and words per minute (WPM), both calculated from the MateCat editing log. Immediately after completing the post-editing task, participants completed the same 15-item Likert-scale questionnaire (post-test) to evaluate their perceptions of post-editing speed and efficiency following the intervention.

The resulting data were analysed using IBM SPSS Statistics (Version 27). Independent-samples t-tests were used to compare PE speed between the experimental and control groups, with Welch's correction applied when the assumption of homogeneity of variance was violated. One-way analysis of covariance (ANCOVA) was conducted to compare post-test questionnaire scores while controlling for participants' pre-test scores.

3.6 Data Collection

3.6.1 Validity

To ensure that the questionnaire accurately measured the intended construct, the items were designed to be clear, concise, and easily understood by the participants. A two-stage validation process was employed: The initial draft of the 15 questionnaire items was submitted to five subject matter experts with recognized expertise in Translation Studies and Translation Technology pedagogy at Iranian universities. These experts were asked to evaluate each item for its relevance to perceived PE speed factors, clarity of wording, potential for ambiguity, and overall coverage of the construct within the context of PE training for students. Feedback from the experts led to minor revisions in the phrasing of four items to improve precision and eliminate potential misunderstandings. The experts collectively agreed that the revised item set adequately covered the intended domain. Designed to complement objective speed data and align with research questions on training effects, the questionnaire measured key subjective constructs: perceived PE time, perceived PE training course, perceived task difficulty, perceived Google Translate quality, self-rated PE speed, and post-task confidence.

3.6.2 Reliability

Cronbach's alpha was employed to evaluate the questionnaire's internal consistency. Individual item statistics were first calculated, followed by the computation of the overall alpha using SPSS. The observed alpha value is 0.82. Although, in theory, refining the instrument by removing items that artificially inflate the reliability of specific dimensions could lead to greater convergence in Cronbach's alpha values, item-level analysis indicated that removing any individual item would reduce the overall reliability coefficient. This suggests that all items contribute to the construct being measured and that none artificially inflate the alpha. These findings, coupled with the calculated Cronbach's alpha of 0.82 for the entire questionnaire, support the instrument's internal consistency and its suitability for research purposes.

Cronbach's Alpha		0.82
Item	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
1	0.615	0.737
2	0.672	0.710
3	0.572	0.759
4	0.696	0.786
5	0.563	0.765
6	0.514	0.837
7	0.667	0.798
8	0.705	0.786
6	0.673	0.796
10	0.652	0.80
11	0.64	0.79
12	0.758	0.747
13	0.67	0.787
14	0.681	0.793
15	0.598	0.821

Table 1. Cronbach's alpha coefficient for the questionnaire

4 Data Analysis

Data analysis was performed using IBM SPSS Statistics V.27. Both descriptive and inferential statistics were employed to examine the effects of the post-editing training program on participants' subjective perceptions of post-editing speed. The significance level for all inferential tests was set at $\alpha = .05$.

4.1 Descriptive Analysis

Before analyzing data statistically, we explore its characteristics through descriptive statistics. Table 2 summarizes and visualizes these statistics for the PE speed variable, separated by hypothesis and analyzed for both pre-test and post-test stages in the two study groups.

Group		Descriptive Statistics					Paired Samples Test
		N Statistic	Mean Statistic	Std. Deviation Statistic	Skewness Statistic	Kurtosis Statistic	
Experimental	post-editing speed. Pre-Test	10	2.63	1.40	-0.17	-1.22	0.000
	post-editing speed. Post-Test	10	3.52	1.42	-0.82	-0.15	
Control	post-editing speed. Pre-Test	10	3.22	1.73	-0.21	-1.21	0.74
	post-editing speed. Post-Test	10	3.24	1.57	-0.01	-1.14	

Table 2. Post-editing speed variable descriptive indicators

Table 2 summarizes the sample sizes, means, and standard deviations of the post-editing speed variables for both groups in the pre- and post-test phases. As shown, the experimental group's mean post-editing speed score increased from 2.63 in the pre-test to 3.52 in the post-test, while the control group's mean increased slightly from 3.22 to 3.24. Notably, both groups experienced an increase in average post-processing speed in the post-test compared to the pre-test. Further analysis using the paired-samples t-test shows a statistically significant difference ($p < 0.05$) in the mean post-editing speed score between the pre- and post-test for the experimental group. This indicates a significant improvement in post-processing speed after the procedure. Conversely, the control group's p-value of 0.74, which is above the 0.05 significance level, indicates that the average post-editing speed score

did not differ significantly between the pre-test and post-test, suggesting that the intervention had no measurable impact on post-processing speed.

4.2 Inferential Analysis

To examine the effect of the PE training program on participants' perceptions of PE speed and efficiency, an ANCOVA was conducted. ANCOVA extends analysis of variance (ANOVA) by comparing group differences on a dependent variable while statistically controlling for the effects of one or more covariates. In the present study, participants' pre-test questionnaire scores were treated as the covariate to control for baseline differences between the experimental and control groups when comparing post-test questionnaire scores. Before conducting ANCOVA, its underlying assumptions were examined to ensure the appropriateness of the analysis. Specifically, the assumptions of normality, homogeneity of variances, and homogeneity of regression slopes were evaluated.

The assumption of normality was assessed using the Kolmogorov–Smirnov test. As shown in Table 3, all significance values exceeded the conventional threshold of 0.05, indicating that the null hypothesis of normality could not be rejected. Therefore, the distributions of the questionnaire scores for both the experimental and control groups at the pre-test and post-test stages did not significantly deviate from a normal distribution.

Group	Phase	Sig.	Decision	Interpretation
Experimental	Pre-test	0.200	Fail to reject H_0	Normally distributed
Experimental	Post-test	0.200	Fail to reject H_0	Normally distributed
Control	Pre-test	0.200	Fail to reject H_0	Normally distributed
Control	Post-test	0.200	Fail to reject H_0	Normally distributed

Table 3. Results of the Kolmogorov–Smirnov normality test for questionnaire scores

The assumption of homogeneity of variances was evaluated using Levene's test. As presented in Table 4, the significance values for both the pre-test ($p = .23$) and post-test ($p = .18$) exceeded the significance level of .05, indicating that the variances of the questionnaire scores were statistically equivalent across the experimental and control groups.

Before conducting ANCOVA, the assumption of homogeneity of variances was assessed using Levene's test. This test examines whether the variance of the dependent variable is statistically equivalent across groups, an assumption required for the valid application of ANCOVA. Table 4 presents the results of Levene's test for the pre-test and post-test questionnaire scores.

Phase	Sig.	Decision	Interpretation
Pre-test	.23	Fail to reject H_0	Variances are homogeneous
Post-test	.18	Fail to reject H_0	Variances are homogeneous

Table 4. Levene's test of homogeneity of variances for questionnaire scores

The homogeneity of regression slopes, another key assumption of ANCOVA, was assessed by testing the interaction between the covariate (pre-test questionnaire scores) and group membership. The interaction was not statistically significant ($p = .48$), indicating that the relationship between the covariate and the dependent variable did not differ significantly between the experimental and control groups. Therefore, the assumption of homogeneity of regression slopes was satisfied. Overall, the assumption tests demonstrated that the questionnaire data met the key assumptions required for ANCOVA. The distributions were approximately normal, the variances were homogeneous across groups, and the homogeneity of regression slopes assumption was satisfied. Consequently, ANCOVA was considered appropriate for comparing post-test questionnaire scores while controlling for participants' pre-test scores.

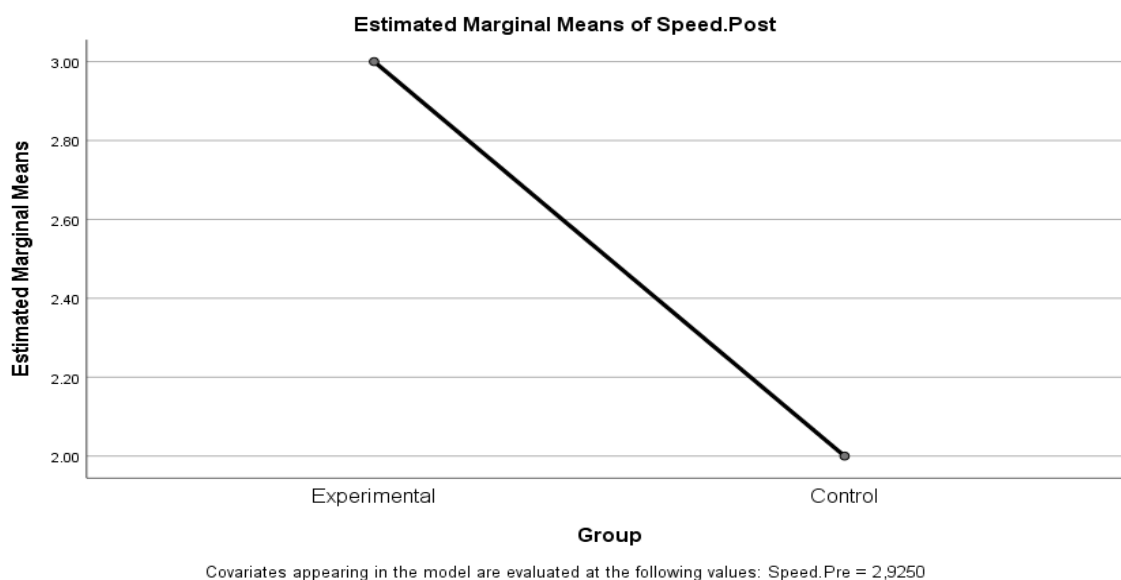
Tests of Between-Subjects Effects					
Dependent Variable: Development of critical thinking and problem-solving abilities					
Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	40.09a	2	20.05	87.91	.000
Intercept	2.29	1	2.29	10.03	.006
Speed.Pre	39.72	1	39.72	174.16	.000
Group	3.18	1	3.18	13.94	.002
Error	3.88	17	.23		
Total	272.79	20			
Corrected Total	43.97	19			

a. R Squared = .912 (Adjusted R Squared = .901)

Table 5. Results of ANCOVA for the post-editing speed variable

Table 5 presents the results of ANCOVA conducted to examine the effect of the PE training program on participants' post-editing speed, while controlling for pre-test speed scores. The overall ANCOVA model was statistically significant, $F(2, 17) = 87.91$, $p < .001$, explaining 91.2% of the variance in post-test speed scores ($R^2 = .912$; Adjusted $R^2 = .901$), indicating that the model provided an excellent fit to the data.

The covariate, Speed.Pre, had a statistically significant effect on post-test speed, $F(1, 17) = 174.16$, $p < .001$, demonstrating that participants' baseline speed scores significantly predicted their post-test performance. This result confirms the appropriateness of including the pre-test measure as a covariate in the ANCOVA model. After controlling for pre-test differences, the effect of Group was also statistically significant, $F(1, 17) = 13.94$, $p = .002$. This finding indicates that the adjusted post-test speed scores differed significantly between the experimental and control groups. Descriptive statistics further illustrate this difference. As shown in Table 5, the control group's mean speed score changed only slightly from 3.22 at the pre-test to 3.24 at the post-test. In contrast, the experimental group's mean speed score increased from 2.63 before the intervention to 3.52 after completing the PE training program.



Graph 1. Trend of changes in the mean of the post-editing speed variable in the post-test stage after removing the effect of the pre-test

As shown in Graph 1, the estimated marginal mean for PE speed in the post-test is higher in the experimental group compared to the control group. This suggests that, on average, the experimental group achieved higher post-test PE speed than the control group after controlling for pre-test PE speed. The graph clearly shows statistically significant improvements in the post-editing speed variable following the PE training program, with a 29% increase from 2.85 in the pre-test to 3.68 in the post-test. These findings suggest the training program had a statistically significant impact on enhancing post-editing speed.

4.3. An Independent Sample t-Test Result

An independent samples t-test was conducted to evaluate potential differences in PE time (in minutes) between the experimental and control conditions. Details of the raw scores, descriptive statistics, inferential results, and data distribution are presented in Tables 6 and 7.

Participant ID	Group	Time Taken (minutes)
P01	Experimental	42
P02	Experimental	38
P03	Experimental	55
P04	Experimental	45
P05	Experimental	48
P06	Experimental	35
P07	Experimental	40
P08	Experimental	50
P09	Experimental	43
P10	Experimental	46
P11	Control	55
P12	Control	62
P13	Control	78
P14	Control	65
P15	Control	71
P16	Control	85
P17	Control	60
P18	Control	75
P19	Control	68
P20	Control	70

Table 6. Post-Editing Time

Group	N	Mean (M)	Std. Deviation (SD)
Experimental	10	44.20	5.92
Control	10	68.90	8.95

Table 7. Descriptive Statistics and Independent Samples T-Test Results for Post-Editing Time (minutes)

Statistic	Value
Mean Difference	-24.70
Welch's t-value	-7.28
Degrees of Freedom (df)	15.61
p-value (2-tailed)	< .001
Cohen's d	3.25

Table 8. Test Statistics

An independent samples t-test was conducted to compare the mean PE times between the experimental and control groups. The results indicated a statistically significant difference, $t(15.61) = -7.28$, $p < .001$. The experimental group ($M = 44.20$ min, $SD = 5.92$) had significantly lower mean post-editing time than the control group ($M = 68.90$ min, $SD = 8.95$). The mean difference between the groups was -24.70 minutes. The effect size was very large (Cohen's $d = 3.25$), indicating a substantial difference in PE speed between the conditions. These findings provide strong evidence that the experimental condition led to significantly faster PE performance compared to the control condition.

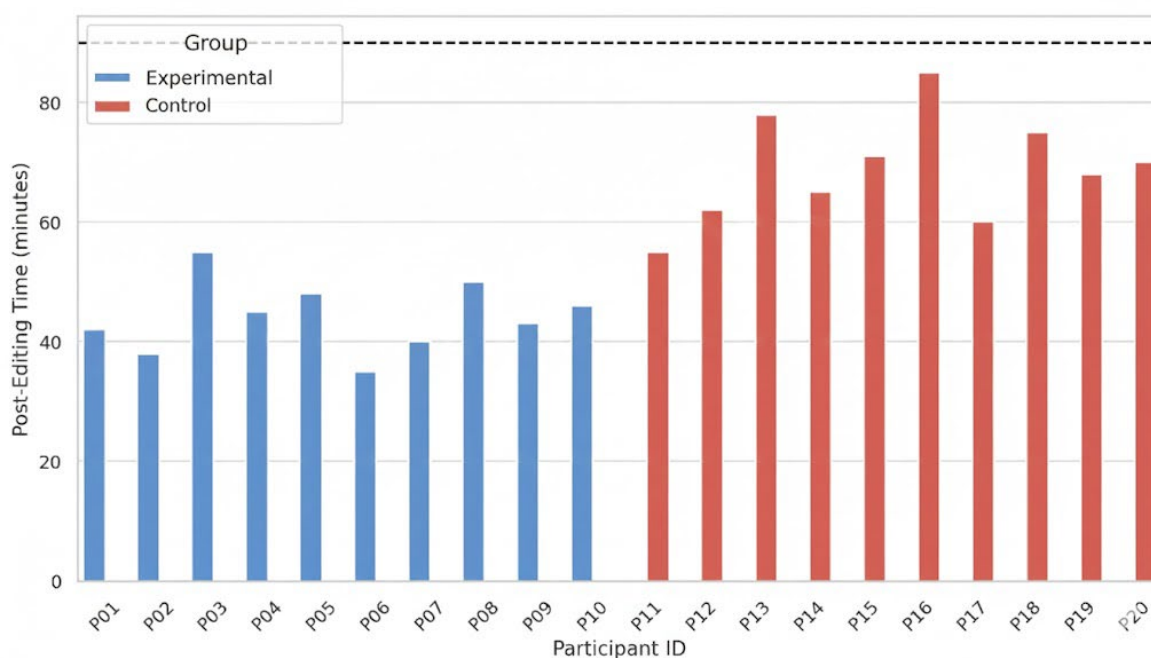


Figure 1. The Distribution of Post-Editing Times

Figure 1 illustrates the distribution of post-editing times for the Experimental and Control groups, utilizing histograms with overlaid density curves. A dashed black vertical line marks the target allocated post-editing time of 90 minutes. The Experimental group's PE times range from 35 to 55 minutes. The distribution is compact, with the highest frequencies ($n = 3$) occurring equally in the 40–45-minute and 45–50-minute bins, both of which are well below the 90-minute target line. Conversely, the Control group's times range from 55 to 85 minutes, indicating a broader, more dispersed distribution, with frequencies of $n = 2$ across the 60–65-, 65–70-, 70–75-, and 75–80-minute intervals. Even the longest control times remain below the 90-minute benchmark, but they are substantially slower than the experimental group's performance. A clear separation between the two distributions is evident, showing minimal overlap. The overlaid density curves emphasize this visual distinction, showing a pronounced leftward shift and a higher peak in the experimental group's distribution relative to the control groups. This visualization graphically supports the statistically significant difference found between the groups ($t(15.61) = -7.28$, $p < .001$, $d = 3.25$), reinforcing the conclusion that the experimental condition substantially reduced post-editing time relative to the 90-minute target.

5 Discussion

This study examined whether a guideline-based PE training programme improves post-editing speed among undergraduate translation students working with the English–Persian language pair. The findings provide strong support for this hypothesis. Participants who completed the PE training

completed the post-editing task significantly faster than those in the control group ($M = 44.20$ vs. 68.90 minutes, $p < .001$, $d = 3.25$) and achieved significantly higher words-per-minute (WPM) rates ($M = 9.20$ vs. 5.89 , $p < .001$, $d = 3.19$). Participants who received the training also reported significantly more positive perceptions of their post-editing speed and efficiency after controlling for baseline scores, $F(1, 17) = 13.94$, $p = .002$. Collectively, these findings suggest that PE speed is not merely a consequence of exposure to MT output but also reflects a component of translation competence that can be enhanced through guideline-based instruction. From the perspective of translation competence, the pattern of results is consistent with a procedural-learning account, whereby repeated practice automates routine editing decisions, reduces cognitive effort, and promotes more efficient post-editing workflows.

These findings align with previous studies showing that PE training improves post-editing productivity and efficiency (Blagodarna, 2018; Samman, 2022; Toledo Báez, 2018) and with research demonstrating that post-editing is generally faster than translation from scratch, although productivity gains depend on translator experience, MT quality, text type, and translation direction (Daems et al., 2017; Jia et al., 2019; Läubli et al., 2019). They also lend empirical support to calls for integrating PE more systematically into translator education in response to the growing adoption of MT in professional translation workflows (Guerberof Arenas & Moorkens, 2019; Kenny, 2022; O'Brien, 2021). By providing experimental evidence from undergraduate translators working with the English–Persian language pair, the present study extends the PE literature to an underexplored linguistic and educational context.

Within the Iranian PE literature, the present study complements the findings of Arjani and Jamshidiha (2024), who found that novice English–Persian translators post-edited more quickly and produced higher-quality translations than when translating from scratch, and Asadi et al. (2023), who demonstrated that improvements in NMT quality reduced post-editing effort for domain-specific terminology. Whereas these studies focused primarily on MT output quality or the PE task itself, the present study indicates that guideline-based instruction is independently associated with faster post-editing, even when the MT system and post-editing task are held constant. This distinction is theoretically important because it suggests that productivity gains arise not only from advances in MT but also from the systematic development of translators' procedural expertise. The findings therefore reinforce competence-based perspectives on PE, which conceptualise post-editing performance as the combined outcome of technological resources and acquired professional competence.

A plausible explanation for these improvements lies in the design of the training programme. Across sixteen sessions, participants repeatedly analysed MT errors, applied the TAUS (2017) post-editing guidelines, monitored their own productivity, and completed authentic PE tasks using the same CAT environment employed during the final assessment. Repeated practice under consistent working conditions likely reduced the cognitive effort associated with error detection, revision, and tool navigation, enabling participants to make editing decisions more rapidly and consistently. This interpretation is consistent with process-oriented research demonstrating that post-editing effort varies according to translator experience, MT quality, and text complexity (Jia & Sun, 2023; Jia et al., 2019; Koglin, 2015) and with de Almeida and O'Brien's (2010) argument that effective post-editing depends on targeted training and experience rather than exposure to MT alone. Notably, the objective improvements in completion time and words per minute (WPM) exceeded participants' self-reported improvements, suggesting that procedural expertise may develop before learners fully recognise gains in their own performance. This finding highlights the importance of evaluating PE training using objective performance indicators alongside self-report measures, as learners' perceptions alone may underestimate the extent of competence development.

Several limitations should be considered when interpreting these findings. The modest sample size, the use of a single financial-domain text, one MT system, one CAT environment, and an exclusively undergraduate sample limit the generalisability of the findings. Furthermore, because training and assessment were conducted within the same instructional and technological environment, the observed improvements are best interpreted as gains in procedural fluency within that context rather than as evidence of broadly transferable PE competence. Future research should therefore examine PE training across larger and more diverse samples, additional language pairs, text genres, MT

systems, and professional translation settings while incorporating delayed post-tests and process-oriented methods, such as keystroke logging, eye-tracking, and think-aloud protocols, to clarify the cognitive mechanisms through which PE training enhances post-editing speed and procedural expertise.

6 Conclusion

This study examined the effectiveness of a guideline-based PE training program in improving the PE speed of Iranian undergraduate translation students working with English–Persian financial texts. The findings demonstrate that PE training significantly improved objective post-editing performance, as reflected in faster task completion and higher words-per-minute (WPM) rates, while also enhancing participants' perceptions of their post-editing speed and efficiency. Together, these findings indicate that PE speed is a trainable component of translation competence that can be developed through systematic, guideline-based instruction.

The study contributes to the growing PE literature by providing empirical evidence from the under-researched English–Persian language pair and highlights the pedagogical value of integrating structured PE training into undergraduate translation programmes. Although the findings should be interpreted in light of the limitations discussed in the preceding section, future research should examine the long-term retention and transferability of PE skills across different text genres, MT systems, language pairs, and learner populations.

Acknowledgment

We gratefully acknowledge the invaluable guidance of our supervisor and the contributions of all participants and supporters in completing this research paper.

References

- Al Sammarraie, H. S. M. (2025). Challenges and strategies in post-editing English into Arabic Neural Machine Translations of movie subtitles. *Journal of Modern Languages*, 35(1), 139-164. <https://doi.org/10.22452/jml.vol35no1.7>
- Al Sammarraie, H. S. M., & Ghassemiazghandi, M. (2024). Enhancing post-editing machine translation skills among Iraqi undergraduate students. *Arab World English Journal for Translation & Literar Studies*, 8(2), 165–182. <https://doi.org/10.24093/awejtls/vol8no2.12>
- Ali, A. F., & Alhebshi, S. H. S. (2026). The Use of AI in Translating Economic and Financial Texts. In D. Chen, A. Sun & D. Li (Eds.), *Translating Finance and Economics: Theory, Practice, Training and Technology* (pp. 99-128). Springer Nature Singapore.
- Allen, J. (2001). Postediting: An integrated part of a translation software program. *Language International*, 13(2), 26–29.
- Alvarez-Vidal, S., & Oliver, A. (2023). Assessing MT with measures of PE effort. *Ampersand*, 11, Article 100125. <https://doi.org/10.1016/j.amper.2023.100125>
- Aranberri, N. (2017). What do professional translators do when post-editing for the first time? First insight into the Spanish-Basque language pair. *HERMES – Journal of Language and Communication in Business*, (56), 89–110. <https://doi.org/10.7146/hjlc.v0i56.97235>
- Arjani, S. H., & Jamshidiha, M. (2024). A comparison of the quality and speed of post-edited translation and human-initiated translation by novice translators. *Iranian Journal of Translation Studies*, 21(84), 54-73. <https://journal.translationstudies.ir/ts/article/view/1074>
- Asadi, V., Shokpour, N., O'Neill, S., Dann, C., & Wang, J. (2023). *Neural machine translation and post-editing: Improving translation of domain specific English terms for a low resource language* [Preprint]. Research Square. <https://doi.org/10.21203/rs.3.rs-3145838/v1>
- Bar-Hillel, Y. (1960). The present status of automatic translation of languages. *Advances in Computers*, 1, 91–163. [https://doi.org/10.1016/S0065-2458\(08\)60607-5](https://doi.org/10.1016/S0065-2458(08)60607-5)
- Bastan, M., Khadivi, S., & Homayounpour, M. M. (2017). Neural machine translation on scarce-resource condition: A case study on Persian-English. In *2017 Iranian Conference on Electrical Engineering (ICEE)* (pp. 1485-1490). IEEE. <https://doi.org/10.1109/iranianee.2017.7985278>
- Blagodarna, O. (2018). *Enhancement of post-editing performance: Introducing machine translation post-editing in translator training* [Doctoral dissertation, Universitat Autònoma de Barcelona]. UAB. <http://hdl.handle.net/10803/666847>
- Blagodarna, O. (2020). Acquisition of post-editing competence: Training proposal and its outcomes. In C. Szabó, M. Fischer, & R. Besznyák (Eds.), *Fit-for-market translator and interpreter training in a digital age* (pp. 85–102). Vernon Press.

- Borucinsky, M., Kegalj, J., & Vukalović, N. (2022). A review of computer-assisted translation (CAT) tools used by Croatian translators. *Hieronymus: Journal of Translation Studies and Terminology*, (9), 83–105. <https://doi.org/10.17234/Hieronymus.9.4>
- Bowker, L. (2019). Fit-for-purpose translation. In M. O'Hagan (Ed.), *The Routledge handbook of translation and technology* (1st ed., pp. 453–468). Routledge.
- Carl, M., Dragsted, B., Elming, J., Hardt, D., & Jakobsen, A. L. (2011). The process of post-editing: A pilot study. *Copenhagen Studies in Language*, 41(1), 131–142.
- Çetiner, C., & İşısağ, K. (2019). Undergraduate-level translation students' attitudes towards machine translation post-editing training. *International Journal of Languages' Education and Teaching*, 7(1), 110–120.
- Costa, F. A., Ferreira, T. C., Pagano, A. S., & Meira, W. (2021). Exploring semantic annotations to measure post-editing quality. In M. Ji & M. P. Oakes (Eds.), *Corpus exploration of lexis and discourse in translation* (pp. 83–104). Routledge.
- de Almeida, G., & O'Brien, S. (2010). Analysing post-editing performance: Correlations with years of translation experience. In *Proceedings of the 14th Annual Conference of the European Association for Machine Translation* (pp. 27–32). European Association for Machine Translation.
- Daems, J., Vandepitte, S., Hartsuiker, R. J., & Macken, L. (2017). Identifying the machine translation error types with the greatest impact on post-editing effort. *Frontiers in Psychology*, 8, 1282. <https://doi.org/10.3389/fpsyg.2017.01282>
- Dede, V., & Antonova-Ünlü, E. (2022). Does a formal post-editing training affect the performance of novice post-editors? An experimental study. *Cankaya University Journal of Humanities and Social Sciences*, 16(2), 131–148. <https://doi.org/10.47777/cankujhss.1106326>
- Devatine, N., & Abraham, L. (2024). *Assessing human editing effort on LLM-generated texts via compression-based edit distance* (arXiv:2412.17321). arXiv. <https://doi.org/10.48550/arXiv.2412.17321>
- Deoghare, S., Kanojia, D., & Bhattacharyya, P. (2024). *Together we can: Multilingual automatic post-editing for low-resource languages* (arXiv:2410.17973). arXiv. <https://doi.org/10.48550/arXiv.2410.17973>
- Doherty, S., & Kenny, D. (2014). The design and evaluation of a statistical machine translation syllabus for translation students. *The Interpreter and Translator Trainer*, 8(2), 295–315. <https://doi.org/10.1080/1750399x.2014.937571>
- Eljazouli, A., & Azmi, N. (2024). Linguistic and terminological complexities in post-editing English-Arabic machine translations. *International Journal of Language and Literary Studies*, 6(3), 16–29. <https://doi.org/10.36892/ijlls.v6i3.1775>
- Emara, N. Y. (2023). Using Machine Translation Error Identification to Improve Translation Students' Post-Editing Skills. *Transcultural Journal of Humanities and Social Sciences*, 4(1), 345–357. <https://doi.org/10.21608/tjhss.2023.289357>
- Garcia, I. (2011). Translating by post-editing: Is it the way forward? *Machine Translation*, 25, 217–237. <https://doi.org/10.1007/s10590-011-9115-8>
- Goldfarb, S. (2024, February 2). Treasury yields rebound after Fed signals higher rates ahead. *The Wall Street Journal*. <https://www.wsj.com/economy/treasury-yields-rebound-after-fed-signals-higher-rates-ahead-b12dd481>
- González-Pastor, D. (2024). La traducción automática y la formación de traductores en España: Perspectivas desde la industria y el ámbito académico. *Mutatis Mutandis: Revista Latinoamericana de Traducción*, 17(1), 117–142.
- Google. (n.d.). *Google Translate* [Software]. Retrieved September 25, 2024, from <https://translate.google.com/>
- Green, S., Heer, J., & Manning, C. D. (2013). The efficacy of human post-editing for language translation. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 439–448). ACM. <https://doi.org/10.1145/2470654.2470718>
- Gu, Q., Wang, S., Ren, X., & Ji, X. (2025). Post-editing of AI-assisted Translation in Translator Education: A Genre-based Classroom Study. *Journal of Historical, Cultural and Social Sciences*, 1(2), 10–22.
- Guerberof Arenas, A., & Moorkens, J. (2019). Machine translation and post-editing training as part of a master's program. *JoSTrans: The Journal of Specialised Translation*, (31), 217–238. <https://doi.org/10.26034/cm.jostrans.2019.184>
- Harto, S., Hamied, F. A., Musthafa, B., & Setyarini, S. (2022). Exploring undergraduate students' experiences in dealing with post-editing of machine translation. *Indonesian Journal of Applied Linguistics*, 11(3), 696–707. <https://doi.org/10.17509/ijal.v11i3.42825>
- Jia, Y., Carl, M., & Wang, X. (2019). How does the post-editing of neural machine translation compare with from-scratch translation? A product and process study. *JoSTrans. The Journal of Specialised Translation*, (31), 60–86. <https://doi.org/10.26034/cm.jostrans.2019.177>
- Jia, Y., & Sun, S. (2023). Man or machine? Comparing the difficulty of human translation versus neural machine translation post-editing. *Perspectives*, 31(5), 950–968. <https://doi.org/10.1080/0907676X.2022.2129028>
- Jovein, S. K., Sharifabad, E. D., & Yazdani, M. (2024). On the relationship between university translator training programs and the translation market requirements: The case of English translation graduates and postgraduates of Imam Reza International University. *Journal of Critical Studies in Language and Literature*, 5(6), 1–13.
- Kenny, D. (2022). Human and machine translation. In D. Kenny (Ed.), *Machine translation for everyone: Empowering users in the age of artificial intelligence* (pp. 23–49). Language Science Press. <https://doi.org/10.5281/zenodo.6759976>
- Khoury, O., Alkhawaja, L., Ghnaim, F., Awwad, S., & Dudeen, H. (2024). Post-editing a Google Translate output: Experienced translators vs. trainees. *Theory and Practice in Language Studies*, 14(6), 1745–1756.

- Kim, H. S., & Cha, Y. (2023). Post-editing of machine translation while reading on English proficiency levels. *Linguistic Research*, 40(3), 517–541.
- Koehn, P. (2009). A process study of computer-aided translation. *Machine Translation*, 23(4), 241–263. <https://doi.org/10.1007/s10590-010-9076-3>
- Koglin, A. (2015). An empirical investigation of the cognitive effort required to post-edit machine-translated metaphors compared to the translation of metaphors. *Translation & Interpreting: The International Journal of Translation and Interpreting Research*, 7(1), 126–141.
- Koponen, M., Salmi, L., & Nikulin, M. (2019). A product and process analysis of post-editor corrections on neural, statistical and rule-based machine translation output. *Machine Translation*, 33(1–2), 61–90. <https://doi.org/10.1007/s10590-019-09228-7>
- Koskinen, K., & Ruokonen, M. (2017). Love letters or hate mail? Translators' technology acceptance in the light of their emotional narratives. In Kenny, D. (ed), *Human issues in translation technology* (pp. 26–42). Routledge.
- Krings, H. P. (2001). *Repairing texts: Empirical investigations of machine translation post-editing processes*. Kent State University Press.
- Kriston, A. (2017). Machine Translation in Translating Business Texts. *Professional Communication and Translation Studies*, (10), 126–136.
- Läubli, S., Amrhein, C., Düggelin, P., Gonzalez, B., Zwahlen, A., & Volk, M. (2019). Post-editing productivity with neural machine translation: An empirical assessment of speed and quality in the banking and finance domain. In *Proceedings of Machine Translation Summit XVII: Research Track* (pp. 267–272). European Association for Machine Translation.
- MateCat. (n.d.). *MateCat* [Software]. Retrieved September 25, 2024, from <https://www.matecat.com>
- Moorkens, J., & O'Brien, S. (2015). Post-editing evaluations: Trade-offs between novice and professional participants. In *Proceedings of the 18th Annual Conference of the European Association for Machine Translation* (75–81). European Association for Machine Translation.
- Nitzke, J. (2019). *Problem-solving activities in post-editing and translation from scratch: A multi-method study*. Language Science Press.
- O'Brien, S. (2011). Towards predicting post-editing productivity. *Machine Translation*, 25(3), 197–215.
- O'Brien, S. (2021). Post-editing. In L. van Doorslaer & Y. Gambier (eds.), *Handbook of Translation Studies* (pp. 178–184). John Benjamins.
- Ortiz-Boix, C., & Matamala, A. (2017). Assessing the quality of post-edited wildlife documentaries. *Perspectives*, 25(4), 571–593.
- Pérez, C. R. (2024). Re-thinking machine translation post-editing guidelines. *JosTrans. The Journal of Specialised Translation*, (41), 26–47. <https://doi.org/10.26034/cm.jostrans.2024.4696>
- Pilevar, M. T., & Faili, H. (2010). PersianSMT: A first attempt to English-Persian statistical machine translation. In *Proceedings of the 10th International Conference on Statistical Analysis of Textual Data (JADT 2010)* (pp. 1101–1112).
- Plitt, M., & Masselot, F. (2010). A productivity test of statistical machine translation post-editing in a typical localisation context. *Prague Bulletin of Mathematical Linguistics*, (93), 7–16.
- Rakhimova, D., Karibayeva, A., & Turarbek, A. (2024). The task of post-editing machine translation for the low-resource language. *Applied Sciences*, 14(2), Article 486. <https://doi.org/10.3390/app14020486>
- 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.
- Sanayai Meetei, L., Singh, T. D., Bandyopadhyay, S., Vela, M., & van Genabith, J. (2020). English to Manipuri and Mizo post-editing effort and its impact on low-resource machine translation. In *Proceedings of the 17th International Conference on Natural Language Processing (ICON)* (pp. 50–59). NLP Association of India.
- Sánchez Ramos, M. M. (2025). Machine translation post-editing through emotional narratives: A methodological approach. *Translation and Translanguaging in Multilingual Contexts*, 11(1), 31–47.
- Saricaoglu, A., & Dinçer, B. H. (2025). Machine translation post-editing behavior of undergraduate EFL students across two genres. *CALICO Journal*, 42(1), 121–142. <https://doi.org/10.3138/calico-2024-1225>
- Schäfer, F. (2003). MT post-editing: how to shed light on the “unknown task”. In *Experiences at SAP.EAMT Workshop: Improving MT through other language technology tools: resources and tools for building MT*.
- Snover, M., Dorr, B., Schwartz, R., Micciulla, L., & Makhoul, J. (2006). A study of translation edit rate with targeted human annotation. In *Proceedings of the 7th Conference of the Association for Machine Translation in the Americas: Technical Papers* (pp. 223–231). AMTA.
- Stasimioti, M., & Sosoni, V. (2019). Undergraduate translation students' performance and attitude vis-à-vis machine translation and post-editing: Does training play a role? In *Proceedings of the Translating and the Computer Conference* (Vol. 41, pp. 125–136). AsLing.
- TAUS. (2017). *TAUS quality evaluation using an error typology approach*. <https://cdn2.hubspot.net/hubfs/2734675/Reports,%20ebooks/Error%20Typology%20Best%20Practices%20Guidelines.pdf>
- Temnikova, I. P. (2010). Cognitive evaluation approach for a controlled language post-editing experiment. In *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10)*.

- Terribile, S. (2023). Is post-editing really faster than human translation? *Translation Spaces*, 13(2), 171–199. <https://doi.org/10.1075/ts.22044.ter>
- Toral, A., Wieling, M., & Way, A. (2018). Post-editing effort of a novel with statistical and neural machine translation. *Frontiers in Digital Humanities*, 5, Article 9. <https://doi.org/10.3389/fdigh.2018.00009>
- Torrejón, E., & Rico, C. (2002). Controlled translation: a new teaching scenario tailor-made for the translation industry. In *Proceedings of the 6th EAMT Workshop: Teaching Machine Translation* (107-116). EAMT.
- Vashee, K. (2013). Understanding the economics of machine translation. *Translation Spaces*, 2(1), 125–149. <https://doi.org/10.1075/ts.2.07vas>
- Wang, L., & Wang, X. (2021). Building virtual communities of practice in post-editing training: A mixed-method quasi-experimental study. *JosTrans. The Journal of Specialised Translation*, (36), 193-219. <https://doi.org/10.26034/cm.jostrans.2021.041>
- Yamada, M. (2019). The impact of Google neural machine translation on post-editing by student translators. *JosTrans. The Journal of Specialised Translation*, (31), 87–106. <https://doi.org/10.26034/cm.jostrans.2019.178>
- Yngve, V. H. (1954). The machine and the man. *Mechanical Translation and Computational Linguistics*, 1(2), 20–22.
- Zhang, H., & Torres-Hostench, O. (2022). Training in machine translation post-editing for foreign language students. *Language Learning & Technology*, 26(1), 1–17. <https://doi.org/10.64152/10125/73466>