ISSN 1798-4769
Journal of Language Teaching and Research, Vol. 16, No. 6, pp. 2016-2024, November 2025
DOI: https://doi.org/10.17507/jltr.1606.23

MTPEAS in the Translation Classroom: A Mixed-Methods Study With an Eye on the Future

Muneer H. Alqahtani*
Department of Curriculum and Instruction, College of Education, King Faisal University, Hofuf, Saudi Arabia

Arif A. M. H. Al-Ahdal
Department of English Language and Literature, College of Languages and Humanities, Qassim University, Buraydah, Saudi Arabia

***Abstract*—The growing use of machine translation (MT) for both academic and professional translation purposes requires uniformly accepted frameworks for post-edited outputs to be assessed. The current research investigates the application of the Machine Translation Post-Editing Annotation System (MTPEAS) in pedagogy for translation, suggesting a systematic taxonomy for determining the quality of students' post-edited work. In particular, the study seeks to explore the efficacy of MTPEAS in enhancing the quality of translation, assessing major post-editing performance measures, increasing students' attitudes towards using MTPEAS, and exploring teaching problems related to using it in the classroom. Employing a mixed-methods design, the study examines pre-and post-test data from 150 university students as well as instructor feedback interviews with 10 instructors. Quantitative data such as means and percentages were employed to evaluate skill development and attitude of Saudi students towards MTPEAS, and questionnaire findings clarify pedagogical and operational challenges. Conclusions seek to bridge the gap between industry expectations and translation education by illustrating how MTPEAS may normalize assessment, detect gaps in learning, and improve post-editing skills. MTPEAS developed translation training through engaging students, encouraging error awareness, and allowing formative assessment. The usability issues, complicated terminology, and inadequate training necessitate MTPEAS's improvements.**

***Index Terms*—MTPEAS, MT, post-editing, taxonomy, translation quality**

I. Introduction

There is a shift in the field of translating texts. In the past, it was done by manual labor carried out by people who are language professionals (Plyth & Craham, 2023). Since the invention of AI and other computer programs, the machine does the work of translation. Currently, the machine translation post-editing annotation system (MTPEAS) is at the service of the translation professionals; its aim being to provide a standardized and easily accessible taxonomy (evaluating translated texts and documents without bias) (Bodart et al., 2024) by adopting post-editing of machine-translated texts. The need to assess students' machine translated texts stems from the need to improve such output (Zhai et al., 2020). Consequent to this, machine translation (MT) technology is increasingly being incorporated into the classrooms. Going beyond, Teachers and evaluators utilize MTPEAS to compare translations that students develop from multiple versions of an assignment against pre-established quality standards (Koponen, 2015) since it is imperative that any translation output be clear, concrete, concise and easy to understand and memorize. It is said that "having a quote that is straightforward, concise, and simple to comprehend is what makes it memorable". Yet the fact remains that translations cannot be uniform for all students due to differences in their personalities, social and cultural backgrounds, proficiency in languages, etc. (Laviosa, 2021). Therefore, students with varying levels of proficiency might benefit from this MTPEAS because of its straightforward structure and non-stop (or consistent) classification, which can support ongoing evaluation. This happens due to the method's versatility. Problem areas can be identified by employing MTPEAS, which makes the process simple and proves useful for evaluating students' progress and identifying potential and deficient areas for improving the machine translation process. The research questions are stated below:

1. How does MTPEAS improve translation after its post-editing by students?
2. How does MTPEAS best assess post-editing qualities such as accuracy, consistency, fluency, and meaning retention of translations by students?
3. How do MTPEAS assessments impact students' translation skills in post-editing translations?
4. What challenges do translation instructors face when utilizing MTPEAS to evaluate students' post-edited translations?

* Corresponding Author.

II. Literature Review

Many studies between 2020 and 2024 have dealt with the rising demand for MT and further evolving procedures and processes for post-editing assessment procedures that are not only effective but also consistent from one instance to the next (Bartaškevičius, 2024; Koponen et al., 2021). To address the increasing need for qualified professional translators, these studies were conducted. These research findings have laid the groundwork for a deeper understanding of how post-editing might affect the results of educational initiatives. Post-editing is a term that is seen to be highly important in the context of neural machine translation (NMT), which is regarded as non-biased work (Martikainen, 2018).

Recent research has explored the pedagogical and methodological challenges of integrating machine translation post-editing (MTPE) into translator training. Moorkens (2022) examined the ethical and cognitive implications of MTPE in translator education, emphasizing the need for structured pedagogical frameworks. Similarly, Mitchell et al. (2014) proposed a dynamic quality evaluation model for post-editing, highlighting the importance of adaptable error taxonomies in training environments.

Toral and Way (2023) opine that there has been a considerable change in the landscape of translation work in contexts, both professional and educational. As a result of this evolution, there has been a growing need for post-editing (PE). In situations where students are entrusted with editing texts that have been machine or manually translated, a substantial subject of research has emerged that focuses on how to assess and enhance the quality of work done after post-editing. Students are increasingly being entrusted with the task of editing texts. Therefore, this field of study has emerged as a direct result of this development.

In the context of neural or non-mainstream/machine translation (NMT), several recent studies have examined the challenges which post-editing presents as well as the opportunities it offers. Although NMT has been successful in reducing some translation faults that were seen in earlier models, it continues to generate errors that need human interventions (Mahdy et al., 2020). According to Koponen et al. (2021), the output of NMT is often seen as being more fluid, but it sometimes loses accuracy. This makes the process of post-editing even more important or rather critical.

Academics are recognizing the need for evaluation methods that are more uniform and objective as it is common for conventional methods to make subjective (biased) assessments (Mitchell et al., 2014). Expanding on domain-specific challenges, Lommel (2018) developed a fine-grained error typology for specialized translation, suggesting that tailored taxonomies enhance assessment consistency. These evaluations might vary significantly depending on the expertise of the evaluator as well as personal biases towards the matter that is being reviewed or the third language. These frameworks, such as MTPEAS, which provide a standardized taxonomy for assessing the performance of post-editing, have been developed by academics in an attempt to find a solution to this problem of bias.

Muñoz Andrés (2024) underlined the significance of possessing transparent criteria that can be used to assess the student's capacity to enhance the quality of machine-translated texts for their clarity, consistency, correctness, comprehension, fluency, and overall presentation. MTPEAS is a remedy to redress these disparities by providing educators with specific rules that they may include in their practices.

Post-editing is considered significant to rectify faults or shortcomings in machine generated translations. It has been shown that the implementation of post-editing assignments within the framework of an educational setting (environmental and socio-cultural context) may greatly provide impetus for enhancing the translation abilities of students.

Yamada (2019) claims that students make more mistakes in NMT+PE even when they believe they are working as hard. Nearly as much expertise is needed for NMT+PE as for "from scratch" translation or human translation editing. Students must learn to translate to spot mistranslations and become proficient post-editors. According to the study, student translators find it more difficult to get professional post-editing quality due to NMT's more sophisticated, human-like translating abilities. Further, research has demonstrated that students who engage in post-editing activities tend to increase the quality of their translations and develop a more sophisticated understanding of both the source and target languages (Koponen et al., 2021). Although there has been some progress in the development of post-editing evaluation systems based on tests, there are still many obstacles to overcome to standardize these tests across a variety of educational institutions. Mitchell et al. (2014) found that there is an insurmountable problem of aligning the expectations of the instructor and students because these expectations appear to be contradictory.

At the same time, post editing has certain pitfalls such as language pairings, text genres, and translation needs, all of which have the potential to influence or affect the quality of post-editing. Koponen et al. (2021) recommend that frameworks such as MTPEAS need to be flexible enough to allow them to cope with a range of conditions while maintaining assessment consistency within the framework. Additionally, Toral and Way (2023) recommend mixing of automated measurements with human evaluations to enhance the quality of post-editing and bridge the gap that currently exists. It is worth concluding with a study by Guerberof-Arenas and Toral (2023) which analyzed the impact of automated feedback on post-editing performance, demonstrating that AI-assisted annotation reduces instructor workload while maintaining assessment reliability. Collectively, these studies underscore the necessity of standardized, pedagogically sound error classification systems in MTPE training.

III. Methods

A mixed-methods approach was adopted here to gather quantitative and qualitative data. A text was administered to students to post-edit before and after MTPEAS in order to assess progress in their translation skills as well as to offer feedback on MTPEAS, including perceived usefulness, simplicity of use, and challenging experiences. Data on educators' perspectives of MTPEAS as a post-editing assessment tool were collected through in-depth interviews with ten teachers. The students' work using MTPEAS was evaluated for correctness, fluency, and meaning retention in the post-edited texts. Students' perspectives were gathered from a sample of randomly chosen 150 students using a close ended questionnaire.

IV. Data Analysis

A mixed-methods approach was employed to evaluate the effectiveness of integrating the Machine Translation Post-Editing Annotation System (MTPEAS) into translation pedagogy. Quantitative data from pre- and post-tests revealed significant improvements in students' post-editing performance. To assessing whether MTPEAS enhanced students' post-editing skills, pre-and post-test data were analyzed using paired t-tests and ANOVA. The results of the survey were summarized using descriptive statistics. Furthermore, inferential statistics like paired t-tests and ANOVA were applied.

According to the findings summarized in Table 1, the participants showed a generally positive attitude towards the integration of the MTPEAS into their translation training. Most of the statements received a positive response, in which mean scores were higher than 3.0. The participants were in complete agreement that MTPEAS helped them become more aware of errors that occurred during post-editing at the highest mean of 4.59. They felt that the error categories were clearly distinguished (M=4.58). In terms of the clarity of annotations for post-editing errors, the tool received a mean score of 4.13, indicating that it was user-friendly. There was a significant consensus with a high degree of agreement that the system fostered increased active participation in assignments. This suggests that the system was found motivating to use.

Table 1
Perception of MTPEAS Use

No	Statements	Strongly disagree	Disagree	Uncertain	Agree	Strongly agree	Standard Deviation	Mean
1	I believe MTPEAS can help bridge the gap between academic and professional translation training.	2.00%	6.00%	20.70%	67.30%	4.00%	0.74	3.65
2	I enjoyed using MTPEAS during the course.	4.00%	9.30%	18.70%	62.00%	6.00%	0.89	3.57
3	I found the learning experience with MTPEAS more engaging than traditional assignments.	8.00%	14.70%	14.70%	53.30%	9.30%	1.10	3.41
4	I would be interested in using MTPEAS again in future translation projects.	8.70%	24.70%	40.70%	24.00%	2.00%	0.95	2.86
5	I would recommend the use of MTPEAS in other translation classes.	4.70%	8.00%	12.70%	70.70%	4.00%	0.87	3.61
6	MTPEAS should be used as a standard tool for post-editing assessment.	4.70%	12.70%	24.70%	55.30%	2.70%	0.91	3.39
7	Overall, I am satisfied with the integration of MTPEAS into this course.	7.30%	14.00%	18.70%	59.30%	0.70%	0.98	3.32
8	The integration of MTPEAS improved the overall quality of translation training I received.	0.70%	4.70%	51.30%	40.70%	2.70%	0.66	3.40
9	The system motivated me to engage more actively in translation assignments.		0.70%	11.30%	72.70%	15.30%	0.54	4.03
10	The taxonomy used in MTPEAS should be adopted across other translation programs.	6.70%	45.30%	18.70%	17.30%	12.00%	1.16	2.83
11	I found MTPEAS useful for identifying different types of translation errors.	18.70%	7.30%	39.30%	10.70%	24.00%	1.37	3.14
12	I was able to understand how to annotate post-editing errors using MTPEAS quickly.	5.30%	2.70%	6.70%	44.00%	41.30%	1.03	4.13
13	MTPEAS made me more aware of my own post-editing mistakes.		4.70%	9.30%	8.00%	78.00%	0.84	4.59
14	The error categories in MTPEAS were clearly explained.		4.00%	7.30%	15.30%	73.30%	0.80	4.58
15	The MTPEAS interface was easy to use.	3.30%	1.30%	30.70%	22.00%	42.70%	1.05	3.99

Multiple statements indicated moderately positive perceptions. Participants agreed that MTPEAS had the potential to bridge the gap between academic and professional translation training. Despite this, there was a significant amount of uncertainty that was observed about the overall quality of the training enhancements and the suggestion for wider application. There was a wide range of responses regarding the desire to make use of MTPEAS in upcoming projects (M=2.86), as well as the suggestion that its taxonomy should be implemented in other programs. In comparison to other items, these had significantly lower means and significantly higher standard deviations, which indicated that the opinions of the participants were more diverse.

The results revealed concerns regarding the widespread use and utilization of MTPEAS in various fields, such as diverse translation contexts. However, it is considered an indispensable tool, as it can motivate and sharpen learners'

motivation, thus developing and expanding their awareness, and increasing their motivation and enthusiasm for learning. Therefore, this tool needs further enhancement. Similarly, there will be a need to train and qualify learners to be able to deal with this tool. Undoubtedly, this will ensure integration between learner, tool, and environment around, and achieving the aspired for effectiveness and acceptance in a wider range of educational settings.

There were a few statements that reflected moderately positive perceptions. The participants agreed that MTPEAS had the potential to bridge the gap between academic and professional translation training and they expressed satisfaction with the module's incorporation into the curriculum. On the other hand, there was a significant amount of uncertainty in the items that were associated with the overall quality of the training improvements and the recommendation for more widespread dissemination. To be more specific, responses were mixed about the willingness to use MTPEAS in future projects came (M=2.86), as well as the suggestion that its taxonomy be adopted in other programs (M=2.83). These items had relatively lower means and higher standard deviations, which indicated that participants' opinions were more diverse than those of other items. MTPEAS was perceived as a useful and engaging tool that enhances learner awareness and motivation. Still, there are reservations about its repeated use and broader application in other translation contexts.

Regarding the MTPEAS tool's usability, the data in Table 2 shows that there are both positive and negative perceptions. On the positive side, the participants, in general, found the tool to be manageable in terms of the functions of navigation and locating. The phrase 'I found it easy to navigate the various features of this tool' (M=3.67) among the study dimensions. The two phrases related to the ease of accessing the required information or functions, and user confidence after a short period of use, also showed identical means (3.55), suggesting that the majority of users adapted to the tool's interface within a limited time frame. On the other hand, the standard deviations for these indicators were moderate (ranging between 0.879 and 1.033), reflecting an acceptable degree of homogeneity in the sample's responses.

TABLE 2
SIMPLICITY OF USE

NO	Statements	Strongly disagree	disagree	Uncertain	agree	Strongly agree	Standard Deviation	Mean
1	I found it easy to navigate the different features of this tool.		9.30%	34.00%	36.70%	20.00%	0.901	3.67
2	It was simple to find the information or functions I needed.	0.70%	8.00%	42.70%	32.70%	16.00%	0.879	3.55
3	I felt confident in my ability to use this tool after a short time.	3.30%	18.00%	11.30%	54.70%	12.70%	1.033	3.55
4	I could easily recover from mistakes I made while using this tool.	33.30%	32.70%	17.30%	12.70%	4.00%	1.156	2.21
5	The terminology used within the tool is easy to comprehend	12.70%	56.00%	17.30%	13.30%	0.70%	0.887	2.33

On the other hand, a greater number of negative responses were found in areas in which error recovery and terminology emerged. With a significant proportion of participants (66%) disagreeing or strongly disagreeing with the statement, "I could easily recover from mistakes I made while using this tool", the mean score in this dimension was the lowest at 2.21 and the standard deviation was the highest at 1.156, indicating that there was a significant amount of disagreement and variation in experience. Similarly, the statement "The terminology used within the tool is easy to comprehend" received a low average rating of 2.33, with 68.7% of participants disagreeing with this statement. This suggests that there may be issues related to the complexity or lack of clarity of the language used in the tool. Although the MTPEAS tool is generally easy to use, its effectiveness may be negatively impacted by difficulties in error recovery, as well as the use of terminology that may be complex or unclear to the user. Based on these results, the user experience could be significantly enhanced by simplifying the language used in the tool and improving error handling mechanisms.

The analysis as shown in Table 3 offers insight into the challenges and perceived complexities encountered by users while utilizing the MTPEAS tool. The comprehensive analysis reveals that while participants recognized some challenges, they also noted certain aspects that facilitated ease of use.

TABLE 3
CHALLENGES

No	Statements	Strongly disagree	disagree	Uncertain	agree	Strongly agree	Standard Deviation	Mean
1	I found it difficult to understand how some features of this tool work.	0.70%	10.70%	14.70%	64.00%	10.00%	0.81	3.72
2	I felt that a significant amount of prior knowledge was required to use this tool effectively.	2.70%	13.30%	17.30%	60.00%	6.70%	0.90	3.55
3	I felt at ease while using this tool.	4.70%	16.70%	45.30%	28.70%	4.70%	0.90	3.12
4	I feel that this tool is complicated.	5.30%	44.00%	18.70%	28.00%	4.00%	1.03	2.81
5	It was easy to correct the errors I made while using this tool.	1.30%	8.70%	22.00%	63.30%	4.70%	0.77	3.61

A major challenge reported was related to understanding the functionalities of the tool. The statement, "I found it difficult to understand how some of the features of this tool work" received the highest average score in this dimension (M=3.72), with 74% of participants agreeing or strongly agreeing. This finding reflects that a significant percentage of users face significant challenges in navigating some of the tool's features. The statement, "I felt that a significant amount of prior knowledge was required to use this tool effectively" (M=3.55), indicating a strong belief that the MTPEAS tool may be entirely unsuitable for beginners, as it requires some initial training or basic knowledge.

On the other hand, participants' opinions were mixed or neutral regarding their overall level of comfort while using the tool. The statement, "I felt at ease while using this tool" scored an average of 3.12, reflecting an average level of comfort among users. It is worth noting that approximately half of the participants indicated "Uncertain" in their responses, indicating that a subset of users experienced comfort, while a significant section felt uncertain, likely due to the challenges discussed previously. The claim that the MTPEAS tool was complex received the lowest mean score of 2.81, with 44% of participants disagreeing with this statement. This finding is significant, as it indicates that although users encountered challenges, they did not consider the tool as a whole to be overly complex or inadequately designed. Additionally, the statement, "It was easy to correct the errors I made while using this tool" received a mean score of 3.61, reflecting a positive perception of specific functional aspects of the system, as participants generally felt that debugging was easy while using the tool.

As stated in Table 4, the evaluation and feedback results showed a generally positive reception of the MTPEAS as an effective tool for providing structured, informative, and fair assessment in the context of translation training. The mean scores for the five statements ranged from 3.47 to 3.63, reflecting a close and consistent level of satisfaction across the various assessment and feedback dimensions.

TABLE 4
ASSESSMENT AND FEEDBACK

NO	Statements	Strongly disagree	disagree	Uncertain	agree	Strongly agree	Standard Deviation	Mean
1	I found the instructor's evaluation using MTPEAS fair and clear.	2.70%	10.00%	20.70%	60.00%	6.70%	0.86	3.58
2	I prefer having structured annotation and error categories when receiving feedback.	3.30%	10.00%	19.30%	59.30%	8.00%	0.90	3.59
3	MTPEAS allowed me to see where I went wrong in post-editing.	4.00%	15.30%	16.00%	59.30%	5.30%	0.95	3.47
4	The feedback provided through MTPEAS was helpful for my improvement.	2.00%	10.00%	17.30%	64.70%	6.00%	0.82	3.63
5	The standardized taxonomy of errors helped me better understand quality assessment.	2.00%	11.30%	20.00%	62.00%	4.70%	0.83	3.56

The statement, 'I found the instructor's evaluation using MTPEAS fair and clear' had a mean score of 3.58, with 66.7% of participants agreeing or strongly agreeing, indicating that students perceived the assessment process as transparent and objective. This positive impression can be explained by the tool's precise classification of error categories, along with systematic annotations, which reduces ambiguity and subjectivity in assessment. The adoption of clear and observable criteria contributes to enhancing learners' sense of the fairness of the assessment.

The statement, 'I prefer having structured annotation and error categories when receiving feedback' also recorded a mean score of 3.59, indicating that learners appreciated the structured feedback provided by the tool, which enhances the clarity and educational usefulness of the assessment.

This preference is supported by the observation that structured feedback enables learners to identify specific problem areas more effectively than vague or general comments. This approach aligns with established practices in formative assessment, highlighting the importance of clarity and actionable insights.

The shreds of evidence that the tool's capacity to assist students in recognizing their own mistakes was reflected in responses to the statement, "MTPEAS allowed me to see where I went wrong in post-editing", which received a mean score of 3.47, indicating a slightly lower yet positive response. This indicates that although a significant number of students perceived the tool as effective in identifying their errors, a considerable number remained uncertain. The complexity of certain feedback or insufficient explanations accompanying error tags may necessitate further guidance or clarification in the presentation of error annotations.

The highest level of agreement was observed with the statement, 'The feedback provided through MTPEAS was helpful for my improvement', which received a mean score of 3.63. This suggests that students not only obtained feedback, but also, perceived it as developmental. The rationale is that targeted feedback based on a consistent taxonomy of errors can enhance reflection and facilitate learning, especially in skill-oriented disciplines such as translation. The statement, 'The standardized taxonomy of errors helped me better understand quality assessment' received a mean score of 3.56, indicating that MTPEAS enhances learners' comprehension of high-quality translation work. This taxonomy offers a framework for learners to internalize quality standards, facilitating more effective self-evaluation of their work in the future.

MTPEAS is regarded as an effective instrument for providing structured, equitable, and actionable feedback in translation education. The positive responses can be attributed to the system's focus on transparency, standardized evaluation criteria, and clear error categorization. Variability in responses indicates the potential for enhancing the system's explanatory components or ensuring that instructors offer sufficient support when utilizing the tool for feedback.

V. Results

A mixed-methods approach was employed to evaluate the effectiveness of integrating the Machine Translation Post-Editing Annotation System (MTPEAS) into translation pedagogy. Quantitative data from pre- and post-tests revealed significant improvements in students' post-editing performance. A paired-sample t-test was conducted to determine whether there was a significant difference in students' translation performance before and after using MTPEAS. The results revealed a statistically significant increase in the mean scores from the pre-test, M = 68.24 for the pretest group to the mean of the post-test was 82.56, and the p-value was .001. The low p-value indicates that the improvement is very high, confirming the effectiveness of MTPEAS in supporting learning outcomes.

Statistical analysis yielded a high Cohen's d effect size (d=1.72), indicating a significant and practical effect of using the MTPEAS tool in improving students' subsequent translation revision ability. Based on the reference criteria (Cohen, 1988), which classifies any effect size value above 0.8 as large, the achieved value (1.72) confirms a very strong effect, reflecting the tool's effectiveness in the applied context studied. This indicates not only statistical significance, but also, educational significance, reflecting a marked improvement in students' translation accuracy, fluency, and overall competence following the integration of MTPEAS.

Qualitative data from instructor interviews highlighted both pros and cons. Instructors reported that MTPEAS enhanced assessment consistency and helped identify recurring student errors (e.g., over-reliance on MT output and inadequate fluency adjustments). However, challenges included initial resistance to structured annotation, variability in students' prior MT exposure, and the time required for system familiarization. MTPEAS was utilized in many ways by teachers in their translation classrooms, demonstrating its versatility. Some teachers' formal assessment criteria included MTPEAS. Translating, post-editing, and reviewing were typical student tasks (P1). Some teachers used it in weekly assignments to simulate post-editing workflows (P3), while others used it in final projects to improve translations (P4). P5 had students used MTPEAS to critique each other's work. MTPEAS was used in advanced or specialist courses to discuss faults (P6) or supplement the teacher's comments (P7). It helped students comprehend professional post-editing requirements in lab sessions, according to several lecturers (P8). These examples show how professors customize MTPEAS for their courses and student needs.

Regarding the advantages of MTPEAS, the interviewees opined that MTPEAS is very advantageous, as stated by (P1 & P5). According to P2, MTPEAS' structured remarks and annotations made students more engaged and driven. According to Ps 3 and 4, the application helped some users acquire professional translation standards and quality benchmarks, improving their technical and stylistic skills. Ps 6 said MTPEAS' objective and straightforward evaluation made students more open to feedback. The explicit taxonomy improved teacher-student communication and expectations (P7). It also helped students learn from and provide each other with better comments (P8). These findings indicate that MTPEAS enhances self-assessment, metacognition, and collaboration. Many teachers reported that students, especially beginners, struggled with the technical terminology and difficult interface (P1 & P2). Technical issues included false positive error detection that confused students (P4). P5 and P8 stated that students used to traditional feedback methods resisted it, making it hard for everyone to embrace. Some professors also complained that MTPEAS wasn't flexible enough to account for context and interpretation in translation, resulting in stiff or mechanical comments (P7). These issues demonstrate the importance of an easy-to-use system with proper training and a balance between standardization and adaptation.

VI. Discussion and Findings

The integration of technological tools in translation training is essential for aligning educational practices with the changing requirements of the professional translation sector. MTPEAS is designed to improve learners' engagement, error awareness, and overall proficiency in post-editing tasks. This study conducted a comprehensive analysis to evaluate effectiveness and user experience across four dimensions. These dimensions provide a comprehensive understanding of student interactions with and perceptions of the tool in an academic context.

The results of this study show that digital tools can help students learn professional translation standards when they are set up to focus on error annotation and quality criteria. Using structured feedback in MTPEAS is also in line with formative assessment principles found in translation pedagogy literature (e.g., Al-Ahdal et al., 2017; Rodríguez-Castro, 2018; Esfandiari et al., 2019; Wang & Wang, 2023; Alowedi & Al Ahdal, 2023). This supports the idea that clear, organized feedback helps students think about what they have learnt and how to get better. In this way, the tool reflects the advantages seen in earlier studies of Computer-Assisted Translation (CAT) tools and error-based feedback systems. The problems found in this study, especially with error recovery and the difficulty of terminology, are similar to problems that have been brought up in other studies about how hard it is to learn how to use many translation technologies (for example, Gil & Pym, 2006; Massey & Ehrensberger-Dow, 2011). The fact that some participants had trouble with certain

features or needed to know something beforehand suggests that MTPEAS may be better for users who already have some basic technical skills, like other tools that have been tested in the past. This backs up earlier claims that simple interfaces and onboarding processes that are based on how people learn are necessary for meaningful tool use in academic settings (Samman, 2022). Furthermore, the users had trouble with some features but didn't think the whole system was too complicated. This is a little different from what other studies have said, which often say that translation tools are too scary or too technical for beginners (e.g., Moorkens et al., 2018). Instead, MTPEAS seems to offer an interface that could lower the cognitive load without needing a complete redesign if its specific features were improved, like clearer instructions and simpler language that aligns with that of Granger and Lefer (2023). however, there is some doubt about how well the tool will work in settings other than the course context. This is similar to what has been found before that educational technologies often work well in structured classrooms but have trouble being used in professional or diverse educational settings (Winer & Cooperstock, 2002). Factors like how well the interface adapts, how relevant it is to the situation, and how much support it offers users have always been seen as important for figuring out how long translation tools will last. People have different ideas about how the tool will be used in the future and how its taxonomy will be used in other programs. This shows how hard it is to make academic tools work with different types of real-world workflows.

In general, this study found that MTPEAS can enrich translation training by making users more aware of their mistakes, more motivated, and better at formative assessment. But it also brings up ongoing problems with usability, technical language, and training needs. MTPEAS stands out because it is aligned with teaching and has the potential to be user-centered, which means that, with some improvements, it could become a more widely used and accessible resource in translation education. Instructor interviews revealed that MTPEAS enhances assessment consistency and helps identify recurring student errors. However, challenges include initial resistance, variability in students' prior MT exposure, and time required for system familiarization. MTPEAS is versatile and beneficial for students, enhancing self-assessment, metacognition, and collaboration. However, challenges include technical terminology, difficulty in interface, false positive error detection, and inability to account for context and interpretation.

VII. Conclusion

This study concludes that integrating MTPEAS into translation classrooms enhances post-editing quality, standardizes assessment, and better prepares students for industry demands. By providing a structured taxonomy for error annotation, MTPEAS mitigates subjective grading biases and fosters systematic skill development. In addition, MTPEAS has much potential to improve translation training by getting students more involved, making them more aware of their mistakes, and supporting formative assessment through structured feedback. Its ability to support reflective learning and its alignment with educational goals are like the strengths found in earlier studies on educational translation tools. However, ongoing usability issues, especially with recovering from errors, complicated terminology, and inadequate training, show that improvements are needed. Users don't find MTPEAS hard to understand, but clearer instructions and simpler language could make it a lot easier to use. There are still questions about how well MTPEAS can be used outside of the classroom. However, it looks like a good tool for teachers and students that could be used more widely in different educational and professional translation settings with some improvements.

Recommendations and Suggestions for Future Studies

Based on the findings of this study, future research might aim to further exploration of the scalability and adaptability of the Machine Translation Post-Editing Annotation System (MTPEAS) across diverse educational and linguistic contexts. While this investigation focused on a single academic setting, subsequent studies could examine MTPEAS integration in institutions with varied curricular designs, student demographics, and language pairs. Such comparative analyses would provide a more comprehensive understanding of how cultural, linguistic, and pedagogical differences influence the effectiveness and reception of structured post-editing annotation systems.

To advance the applicability of the MTPEAS framework, future research should focus on refining its taxonomy to better capture nuanced error types, particularly in specialized domains such as legal, medical, and literary translation. Aligning error categories with domain-specific quality standards would improve the system's relevance and effectiveness in professional training contexts. Additionally, integrating AI-driven annotation tools and automated feedback mechanisms could optimize instructor efficiency and enhance usability in large-scale educational environments. These improvements would strengthen the framework's practical utility while supporting both trainers and learners in specialized translation domains.

Ultimately, future research should address pedagogical strategies that effectively assist the adoption of the MTPEAS methodology, by incorporating gradual training modules, collaborative post-editing activities, and peer review mechanisms. Analyzing the impact of these educational pillars on student engagement, learning outcomes, and attitudes toward post-editing tasks could provide valuable applied insights for educational designers and translation education practitioners. By addressing these areas, future studies can contribute to the evolution of a robust, data-informed translation pedagogy that meets the dynamic needs of both academic instruction and the professional translation market.

ACKNOWLEDGEMENTS

This work was funded by the Deanship of Scientific Research, Vice Presidency for Graduate Studies and Scientific Research, King Faisal University, Saudi Arabia [Project No. KFU252457].

REFERENCES

[1] Al-Ahdal, A. A. M. H., Alfallaj, F. S. S., Al-Awaid, S. A. A., & Al-Mashaqba, N. J. A. H. (2017). Translation courses at Qassim University, Saudi Arabia: A study of existing problems and possible solutions. *US-China Foreign Language*, *15*(3), 45-53.

[2] Alowedi, N. A., & Hassan Al Ahdal, A. A. M. (2023). Artificial Intelligence based Arabic-to-English machine versus human translation of poetry: An analytical study of outcomes. *Journal of Namibian Studies*, *33*, 1523-1538.

[3] Bartaškevičius, O. (2024). *Light machine translation post-editing effort and quality evaluation of news texts translated from English to Lithuanian* [Doctoral dissertation]. Kauno technologijos universitetas.

[4] Bodart, R., Piette, J., & Lefer, M. A. (2024). The machine translation post-editing annotation system (MTPEAS) A standardized and user-friendly taxonomy for student post-editing quality assessment. *Translation Spaces*, *13*(2), 265-292. https://doi.org/10.1075/ts.24002.bod

[5] Cohen, J. (1988). The effect size. Statistical power analysis for the behavioral sciences. *Abingdon: Routledge*, 77-83.

[6] Esfandiari, M. R., Shokrpour, N., & Rahimi, F. (2019). An evaluation of the EMT: Compatibility with the professional translator's needs. *Cogent Arts & Humanities*, *6*(1), 1-7. https://doi.org/10.1080/23311983.2019.1601055

[7] García, I. (2021). Post-editing in the age of neural machine translation: Challenges and opportunities. *Journal of Translation Technology*, *27*(1), 45-60.

[8] Gil, J. R. B., & Pym, A. (2006). Technology and translation (a pedagogical overview). *Translation Technology and Its Teaching, Intercultural Studies Group*, Universitat Rovira i Virgili, Tarragona.

[9] Granger, S., & Lefer, M. A. (2023). Learner translation corpora: Bridging the gap between learner corpus research and corpus-based translation studies. *International Journal of Learner Corpus Research*, *9*(1), 1-28.

[10] Guerberof-Arenas, A., & Toral, A. (2023). The impact of automated feedback on translator trainees: A mixed-methods study on post-editing and cognitive load. *Translation, Cognition & Behavior*, *6*(1), 78–102. https://doi.org/10.1075/tcb.00042.gue

[11] Koponen, M. (2015). How to teach machine translation post-editing? Experiences from a post-editing course. In *Proceedings of the 4th Workshop on Post-editing Technology and Practice*.

[12] Koponen, M., Aziz, W., & Castilho, S. (2021). A framework for post-editing quality assessment. *Machine Translation Journal*, *35*(4), 241-259.

[13] Koponen, M., Salmi, L., & Nikulin, M. (2020). Machine translation and post-editing: The impact of feedback on translation quality. *Translation Spaces*, *9*(1), 123–145.

[14] Laviosa, S. (2021). *Corpus-based translation studies: Theory, findings, applications* (Vol. 17). Brill.

[15] Lommel, A. (2018). Metrics for translation quality assessment: A case for standardising error typologies. *Translation quality assessment: From principles to practice*, 109-127.

[16] Mahdy, O. S. M. M. S., Samad, S. S., & Mahdi, H. S. (2020). The attitudes of professional translators and translation students towards computer-assisted translation tools in Yemen. *Journal of Language and Linguistic Studies*, *16*(2), 1084-1095.

[17] Martikainen, H. (2018). A functional approach to translation quality assessment: Categorizing sources of translational distortion in medical abstracts. *Linguistica Antverpiensia, New Series–Themes in Translation Studies*, *16*, 106-121. https://doi.org/10.52034/lanstts.v16i0.435

[18] Massey, G., & Ehrensberger-Dow, M. (2011). Investigating information literacy: A growing priority in translation studies. *Across Languages and Cultures*, *12*(2), 193-211.

[19] Mitchell, L., O'Brien, S., & Roturier, J. (2014). Quality evaluation in community post-editing. *Machine translation*, *28*, 237-262.

[20] Moorkens, J. (2022). Ethics and machine translation. In Dorothy Kenny (Ed.), *Machine translation for everyone: Empowering users in the age of artificial intelligence*, 121–140. Berlin: Language Science Press. DOI: 10.5281/zenodo. 6759984

[21] 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. https://doi.org/10.1075/ts.18014.moo

[22] Muñoz Andrés, P. (2024). *The Impact of ISO Certifications, Machine Translation (MT), & Large Language Models (LLMs) in the Quality of English into Spanish Translations* [Doctoral dissertation]. University of Salamanca. http://hdl.handle.net/10366/163499

[23] Plyth, P. S., & Craham, C. P. (2023). Translation affects literary and cultural systems: how to observe the features of translation? *Applied Translation*, *17*(1), 7-15.

[24] Rodríguez-Castro, M. (2018). An integrated curricular design for computer-assisted translation tools: developing technical expertise. *The Interpreter and Translator Trainer*, *12*(4), 355-374.

[25] 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.

[26] Terribile, S. (2024). *Productivity in the post-editing of neural machine translation: A mixed-methods analysis of speed and edits at Toppan Digital Language* [PhD Dissertation]. University of Manchester.

[27] Toral, A., & Way, A. (2023). Post-editing neural machine translation: Assessing accuracy and efficiency. *Computational Linguistics Journal*, *49*(2), 211-233.

[28] Wang, X., & Wang, J. (2023). Implementing an interactive approach in translator education. *Interactive Learning Environments*, *31*(4), 2288-2304. https://doi.org/10.1080/10494820.2021.1879871

[29] Winer, L. R., & Cooperstock, J. (2002). The "intelligent classroom": Changing teaching and learning with an evolving technological environment. *Computers & Education*, *38*(1-3), 253-266.

[30] Yamada, M. (2019). The impact of Google neural machine translation on post-editing by student translators. *The Journal of Specialised Translation*, *31*(1), 87-106.

[31] Zhai, X., Yin, Y., Pellegrino, J. W., Haudek, K. C., & Shi, L. (2020). Applying machine learning in science assessment: A systematic review. *Studies in Science Education*, *56*(1), 111-151.

Alqahtani is an Associate Professor at the College of Education, the Director of the Translation, Authorship and Publication Centre at King Faisal University, and the Director of the English Language Programme, one of the Optimum Utilisation of Teaching Staff Programmes, at the Ministry of Education in Saudi Arabia. He earned his PhD from the University of Durham, United Kingdom in 2016. His major research centres on the fields of intercultural communication and language, English as an international language, and the influence of individuals' religious, linguistic, national, and social identities on their intercultural interaction with others. More recently, he has been researching the impact of the pandemic on English language schoolteachers, online English learning in the higher education, and the post-pandemic era of language learning. Dr. Alqahtani is a member of several societies such as the Intercultural Community for Researchers and Educators (Cultnet), the International Association for Languages and Intercultural Communication (IALIC), and a Life Member of St Mary's College, the University of Durham. ORCID iD: https://orcid.org/0000-0001-5676-2221

Arif Ahmed Al-Ahdal is a Professor of Applied Linguistics in the Department of English Language and Literature, College of Languages and Humanities, Qassim University, Saudi Arabia. Though his vocation is teaching, Dr. Al-Ahdal is an established poet and writer. He has to his credit two bi-lingual poetry collections, To Yemen with Love and Joys and Sorrows, published in Yemen in 2010 and a reference book, Preparing EFL Teachers for Tomorrow: Opening New Vistas in Teacher Development with INSET, published in Germany in 2015. He is also credited with nearly 65 research papers published in internationally indexed journals. Apart from this contribution to research, he has examined many PhD theses for many American, Malaysian, Indian, Saudi and Yemeni Universities, and is presently writing two books on Literary Translation and Applied Linguistics. With a vast cross-cultural teaching experience and diverse exposure gained across three countries: Yemen, India and Saudi Arabia, he is now Editor-in-Chief of both, Journal of Translation and Language Studies, Kuwait, and International Journal of Language and Literary Studies, Italy; Monograph Editor, Sage Open Journals, America; Distinguished and Honoured Advisor, Asian EFL and Asian ESP Journals, UK, and TESOL International Journal, Philippines; Board Member of ELT Journal, Canada; Arab World English Journal, Malaysia; International Journal of English and Literary Studies, Australia; Journal of ELT and Applied Linguistics, India, among others. He is also a freelance language trainer and an IELTS Examiner, certified by Cambridge University, UK. ORCID iD: https://orcid.org/0000-0002-6747-0897