

A MULTIFACETED EXPLORATION OF AI-POWERED GOOGLE TRANSLATE USAGE AND POST-EDITING TRAINING EFFECTS

Yen-Yu Lin

ABSTRACT

This study investigated the impact of using machine translation (MT) and post-editing training on the revision process of thirty intermediate-level EFL learners through a single-group pre-post design. The participants identified specific grammar and semantic errors in texts generated by Google Translate (GT) during the post-editing training. Before and after the training, the students translated their L1 writing into L2 without using GT and then edited their L2 writings by comparing them with the GT translations. Data were collected from various sources, including writing outcomes, screen recordings, perception surveys, and interviews. The results showed that there was a significant difference in word length and word diversity between students' revised texts and L2 texts in both the pretest and posttest. In addition, the error rates in the posttest were much lower than those in the pretest. Moreover, the content similarity rate was found to negatively correlate with the error rates in students' revised texts. The survey revealed that students expressed moderate to high satisfaction with the overall quality of texts generated by GT. The study presents implications for utilizing MT as a support for EFL students' writing along with discussing ways for EFL teachers to incorporate MT into the classroom given its increasing demand.

Key words: machine translation, L2 writing, post-editing, revision

INTRODUCTION

Artificial intelligence and machine learning have improved MT reliability, accuracy, and human-like behavior (Godwin-Jones, 2019). With the advancement of MT technology, an increasing number of studies are exploring the impact of using MT on student writing (e.g., Baker, 2013; Chen et al., 2019; Eriksson, 2021; Lee, 2020, 2022; Nino, 2020; Kol, Schcolnik, and Spector-Cohen, 2018; Tsai, 2019, 2022; Zhang, 2022). While several scholars have found that MT has the potential to assist lower- and intermediate-level learners by reducing cognitive load and improving writing fluency (Briggs, 2018; Garcia & Pena, 2011; Kliffer, 2008), the extent of its benefits for this group remains uncertain. On the other hand, Bahri and Mahadi (2016) and Lee (2020, 2022) found that advanced learners may benefit more from MT due to their stronger language foundations. More empirical research is needed to understand how MT can effectively help lower-level L2 learners improve their writing.

Regarding the importance of instruction on MT use in writing, O'Brien et al. (2018) and O'Brien and Ehrensberger-Dow (2020) have suggested that MT literacy instruction or MT post-editing (MTPE) training should be provided to help learners use MT effectively. Also, Zhang and Torres-Hostench (2022) indicated that students at all levels should be allowed to use MT for correcting minor errors and improving accuracy in their L2 writing. By incorporating the analysis of (mis)translations into educational tasks, students can develop skills in effectively checking MT outputs. As students are accustomed to spotting grammatical errors in language activities, teachers could utilize this method to identify errors in MT output (Lee, 2022). However, research on the benefits of MT post-editing training for second language writing remains scarce.

Moreover, another issue of great concern among scholars is students' (over)reliance on MT (Ahn & Chung, 2020; Liu et al., 2022). While many studies have shown through interviews or questionnaires that both teachers and students worry about the negative effects of over-reliance on MT in the teaching and learning of writing (e.g., Baker, 2013; Eriksson; Nino, 2020), research that actually examines the correlation between students' degree of dependence and their error

rates in revising texts is rare. It is worth noting that although previous studies have revealed improvements in students' writing performance after using MT, such as increased writing length and a reduction in grammatical errors, these studies have neither investigated nor demonstrated the similarity between students' revisions and the machine-translated texts. This makes it difficult to determine the extent of students' reliance on MT and whether its use diminishes students' critical thinking and judgment skills. Analyzing the relationship between text similarity and students' error rates in revising articles, along with observing students' writing processes through screen recording, will help us understand the situation regarding the application of analytical thinking and critical evaluation skills.

To address the research gaps mentioned above, the present study aims to investigate (1) the impact of MT use and post-editing training on intermediate students' writing performance and (2) the relationship between text similarity (students' L2 writing & MT text) and text error rate in students' revised English texts, which has not been discussed in the literature. To accomplish the research objectives, the following six questions are formulated as follows:

1. How do MT use and MTPE affect writing quality? Are there any differences between students' revised texts in the pretest and posttest?
2. Did students rely less on MT output after MTPE? Is there any difference between students' degree of reliance on MT output when revising in the pretest and posttest?
3. Is there a relationship between the content similarity rate and the rate of errors in students' revised texts?
4. For the pretest and posttest, is there any significant difference between the rate of errors in students' L2 texts and revised texts?
5. Is there a significant difference between the rate of errors in students' revised tests in the pretest and posttest?
6. What are the students' perceptions of MTPE and using MT for English narrative writing?

The value of this study lies in exploring the possible changes in the writing quality of intermediate-level students after receiving training and uncovering the extent of students' reliance on MT and its correlation with the error rates of their writing.

LITERATURE REVIEW

In recent years, translation has become a significant focus in language learning as an instructional approach (Wilson & Gonzales Davies, 2017). Cook (2010) suggests a natural link between a learner's first language and their second language in the minds of EFL writers. Beiler and Dewilde (2020) and Kim (2011) note that learners naturally and spontaneously engage in translation when attempting to write in the target language. Furthermore, research has indicated that translation can facilitate the enhancement of L2 writing (Cohen & Brooks-Carson, 2001; Lee, 2020). According to Cohen and Brooks-Carson (2001), translation can provide linguistic support that can improve the development of syntactic complexity and coherence in second language writing by exposing learners to lexical items that are beyond their current competency.

With the belief that translation helps to promote L2 writing skills and reduce cognitive load, numerous studies have looked into the potential of MT as a scaffolding to help students write longer and better essays. Kol, Schcolnik, and Spector-Cohen (2018), for example, discovered that when Israeli EFL students used GT for English Academic Purposes (EAP) writing across various tertiary levels, they produced significantly longer texts and enhanced their vocabulary usage. In 2019, Lee observed that employing MT helped students develop writing strategies and reduce lexical and grammatical errors. Similarly, in Tsai's (2019) study on the impact of GT on EFL drafts for Chinese undergraduate students majoring in English, the findings showed that students' English versions translated from their Chinese texts using GT were notably superior to their self-written English versions, featuring a greater number of words, more advanced vocabulary, and fewer spelling and grammatical errors. Furthermore, Chen, Tsai, and Tsou (2019) discovered that utilizing students'

Chinese with the assistance of GT and referring to the GT-translated versions aided EFL sophomore students in enhancing their English content, improving their credibility while making fewer grammatical or syntactic errors, and presenting a more professional style. Moreover, Tsai in 2022 conducted another study to illustrate that GT is an effective tool for non-English majors in revising impromptu reflective essays after viewing a five-minute excerpt from a film.

While several scholars have suggested that MT could assist low- and intermediate-proficiency students in lowering cognitive load and improving writing fluency (Briggs, 2018; Garcia & Pena, 2011; Kliffer, 2008), others have expressed concerns about its potential impact on lower-level learners. Lee (2020) suggests that students at lower proficiency levels may not fully benefit from MT due to limitations in their language knowledge, confidence, and motivation. Bahri and Mahadi (2016) and Tsai (2019) further emphasize that learners at a low level of proficiency might require guidance from teachers on how to effectively utilize these new language learning technologies. Chung (2020) found that, when asked to post-edit MT output, lower proficiency level students had difficulty in identifying and correcting errors, thus suggesting that MT use and MT-related activities with low and intermediate learners should be carefully conducted with sufficient guidelines. Similarly, in Lee's (2022) study, the impact of students' L2 writing proficiency on their revisions when using MT was investigated. It was observed that students' L2 proficiency and their confidence in writing appeared to affect their ability to identify and correct errors. The higher-level groups were more proactive in making changes compared to the lower-level groups. Furthermore, despite being provided with better options by MT, many global grammatical errors remained unresolved in the lower-level groups, although they did address lexical and local grammatical errors. The study suggested that self-editing requires substantial attentional, cognitive, and linguistic resources (Ferris & Roberts, 2001; Waller & Papi, 2017); lower-level students using MT alone does not contribute to the improvement of L2 writing in classrooms. It is important for teachers to offer clear explanations of language rules, which can help minimize students' uncertainties about linguistic elements.

While the importance of integrating MT literacy instruction and

post-editing training in language learning has been addressed (e.g., O'Brien, Simard, & Goulet, 2018; O'Brien & Ehrenberger-Dow, 2020), only a few studies have attempted to investigate the pedagogical design or teacher interventions in using MT for L2 writing. One of the most common practices in previous research on L2 teachers' use of MT in teaching writing involves introducing post-editing of MT into language classes. These activities are designed for students to practice error correction and observe differences across languages, thereby enhancing their understanding of L2 grammar and vocabulary. Enkin and Mejias-Bikandi (2016) introduced post-editing exercises specifically focusing on three types of Spanish clause structures: nominal complements, relative clauses, and adverbial clauses. They utilized faulty online translator output in a Spanish grammar course to raise students' linguistic awareness of second language grammar as well as differences between grammatical structures in the first and second language. Valijarvi and Tarsoly (2019) also explored integrating MT into the instruction of Finnish and Hungarian at various proficiency levels. In their approach, students engaged in post-editing activities at both the phrase and text levels. At the phrase level, students analyzed error-prone phrases generated by machines to identify common error patterns. At the text level, students worked with text samples from diverse genres, evaluating deficiencies in MT outputs, including genre suitability, information coherence, reference tracking, and overall cohesion. Students reported finding the post-editing exercises enjoyable, particularly when comparing translations provided by the instructor with those generated by GT. Additionally, by participating in the exercise, students became aware of the common mistakes GT makes, thereby reducing the possibility of making fundamental mistakes. Zhang (2022) conducted a study that focused on developing MT post-editing training for Chinese students learning Spanish. The training materials aimed to help learners identify six common mistakes found in MT raw output, including accuracy, word order, official name, preposition, omission, and formal style. Results indicated that the experimental group who underwent this training program demonstrated significant improvements in the post-test. Moreover, they were able to complete the posttest more quickly with fewer pauses and showed more

effective editing skills. Focusing on English as the target language, Shin & Chon (2023) investigated how L2 learners from South Korea employed post-editing strategies to address lexical and grammatical errors in machine-translated texts. The errors in machine translation (MT) output primarily consisted of mistranslations, missing words, ungrammatical sentences, and extraneous words. The error categorization scheme was adapted from previous research on error analysis (e.g., Costa et al., 2015; Lee & Briggs, 2021; Moorkens, 2018). Ungrammaticality, a broad category, encompassed errors related to verb tense, articles, sentence fragments, missing prepositions, incorrect verb forms, incorrect word forms, misplaced adverbs, and word order. The findings revealed that higher proficiency learners perform better in error detection and the application of PE strategies. It is suggested that lower-level L2 learners would benefit from training in using MT, specifically in detecting MT errors. In addition, Nino (2008) examined how MT can be used in foreign language classes. In this class, students learned about MT, its advantages and limitations, the MT process, and what kinds of errors they can anticipate. The result showed that post-editing practice triggers advanced learners' awareness of form and negotiation of meaning, enhancing grammatical and lexical accuracy. O'Neill's study (2016) exemplified the importance of teacher interventions and guidance in using MT in FL contexts. His study examined the effectiveness of using MT among three groups: the non-MT group, the MT group without prior training, and the MT group with prior training. O'Neill discovered that the MT group with prior training had better outcomes than the other two groups in the writing task. They scored the highest on the experimental writing tasks. The researcher argued for training students in the responsible use of translators for second language writing. These findings suggest that practicing the detection and correction of specific types of MT errors can be highly beneficial.

The previous studies reviewed above supported the idea that creating and evaluating specific exercises for MT post-editing (MTPE) in language courses is more effective than having students correct a randomly generated MT text with various errors. Moreover, it has been mentioned that, for students with limited language skills,

depending only on MT without proper direction can lead to misunderstandings or an overreliance on the tools (Nino, 2020). MT post-editing activities must be structurally planned to scaffold these learners in areas that need the most attention. More importantly, as the advantages of MT usage for lower-proficiency learners differ from those of advanced learners, practitioners should take into account the diverse needs of learners at different proficiency levels and tailor MT activities accordingly. So far, only a limited number of studies have focused on whether post-editing training influences the ability and awareness of students at a particular level to identify and rectify errors in their own writing. The impact of such training on students' error detection remains an area that requires further investigation. In addition, as reviewed above, previous studies on MT post-editing training do not focus on Chinese as the source language and English as the target language. There is little discussion about Chinese learners of English and common error types of MT output from Chinese to English. What's more, while previous research emphasizes that students tend to rely on MT (Can, 2023; Jolley & Maimone, 2015; Lee, 2022), there is a noticeable lack of studies that compare the similarity between machine-translated texts and students' revised texts. These comparisons are crucial for understanding the extent to which students depend on MT.

In light of what is mentioned above, the present study investigates whether MT use and post-editing training can help intermediate-level EFL students improve their English narrative writing. It looks specifically at assessing the effectiveness of post-editing (PE) training on narrative writing produced by Chinese learners of English (B1 level). The emphasis on narrative writing stems from research suggesting that "narrative development is foundational for other genres, such as expository and argumentative writing (Grennera et al., 2020, p. 694)." It has been demonstrated that various genres place varying levels of cognitive demands on students, with narrative posing the least cognitive load (Genung, 1900; Bain, 1967; Weigle, 2002). According to Alrajhi's analysis of MT narrative texts, MT's output has higher quality than texts produced by intermediate-level learners and has much more advanced vocabulary and function words (2023). Moreover, Chung & Ahn (2022), in their study regarding the

use of MT in students' narrative writing and argumentative writing, found that much more advanced vocabulary appeared in narrative MT texts and using MT could enhance the accuracy of narrative writing in certain aspects, particularly in checking grammar and sentence structure. These results demonstrate that MT can be a valuable tool to help students identify and correct grammatical errors and expand the vocabulary in their narrative writing.

METHODOLOGY

Participants

The participants of the study comprised 30 Chinese students at a national science and technology university in central Taiwan. They came from the College of Engineering and were enrolled in a freshman English course. Student English proficiency was assessed using the College Student English Proficiency Test (CSEPT) with a total score of 240. This test was developed by the Language Training and Testing Center in Taiwan (LTTC). The mean score of the participants is 193, indicating a proficiency level of B1 according to the Common European Framework of Reference for Languages (CEFR) model.

Instruments

The Background Questionnaire

The background questionnaire obtained participants' demographic information, including their degree major and previous English achievements. It also elicited detail in two areas: (1) participants' prior experience of English writing and (2) their experience of using online tools to facilitate English writing.

Evaluation Questionnaire

The evaluation questionnaire's design focused on gathering feedback from the participants concerning their perception of MT output quality, the support of MT in revising self-written English texts,

and the effect of the post-editing training on students' writing processes. Questions in Alrajhi (2023) were used as a basis for developing the question items. The questionnaire contains two sections. The first section deals with the evaluation of GT output quality. GT's quality was examined through items on grammatical accuracy, vocabulary choice, content accuracy, context relevance, and general quality. The second section focuses on student perceptions of the usefulness of MT texts in enhancing their knowledge of English vocabulary, grammar, and sentence structures. The questionnaire was developed based on a five-point Likert scale (ranging from *strongly agree* to *strongly disagree*). The open question at the end of the second section allows students to express thoughts about their gains and changes after training and suggestions on future training design.

Interviews

Seven days after students performed the writing tasks, stimulus recall, and semi-structured interviews were conducted in Chinese in order to collect further clarification of participants' decisions, challenges faced, specific revision moves, and overall perceptions related to using GT during the writing process. Whenever the researcher felt that more on-site information was needed, she paused the stimuli. Following the interviews, the transcripts were analyzed to identify key themes and categories. The main themes that emerged from the analysis include perceptions of GT's usefulness, strategies for using GT, and challenges encountered. The identified themes were validated by cross-referencing them with student texts and screen recordings. By doing so, the researcher was able to gain a comprehensive understanding of how participants perceived GT and utilized it to enhance their writing performance.

Research Procedure

We chose a pretest-posttest design before and after MT post-editing training. The experiment was carried out weekly for a period of three weeks. In the first week, participants filled out a background questionnaire and took a pretest. In the following week, a 90-minute post-editing training was arranged. In the third week, the students took

a post-test. All students wrote using on Google Docs. For the pretest, students were assigned a writing task to describe a time when they experienced failure and hardship. The students had 15 minutes to write a paragraph of 200 words in Chinese (L1) on their own. After finishing the writing task in Chinese, the students had 20 minutes to translate the Chinese passage into English (L2 text). To ensure that participants relied solely on their own knowledge, additional resources such as online dictionaries and paper dictionaries were not allowed for referencing. This restriction was implemented to measure students' basic writing skills without external influences, enabling a more accurate comparison between the original and revised texts to isolate the impact of machine translation on writing quality. Afterwards, students pasted the Chinese text into GT to generate the English version of the paragraphs. GT, the most widely utilized MT tool among students (Tsai, 2019), was employed in this study because it has been greatly improved in terms of accuracy and fluency (Sun, 2017). The students had to spend 20 minutes comparing their L2 texts with the MT outputs, identifying differences, spotting errors, and revising their L2 texts. The posttest included the same procedure as the pretest session. Students wrote a paragraph to describe an event in the past that shaped his/her present self. The pretest and posttest were captured by a screen recorder named Movavi. Finally, an evaluation questionnaire using the Google Forms platform was administered to the participants.

Seven days after completing the writing task in the posttest session, 15 participants were invited to join in one-to-one stimulus recall sessions and semi-structured interviews. Respondents were selected based on the comparison of content similarity rates in their pretest and posttest L2 texts with GT texts. They were from three groups, with five people in each group. For the first group of respondents, the similarity rate between their pretest and posttest revised texts and the GT texts was higher than 60%. The second group of participants showed a similarity of less than 40% between their pretest and posttest L2 texts and the GT texts. The third group of participants had pretest revised texts that were more than 60% similar to GT texts, but posttest revised texts were less than 40% similar.

PE Training Session Design

To develop the PE training materials, ten Chinese narratives focused on personal experiences were selected from the websites of the Taoyuan Department of Education and Bai Yun Piao Piao, in which there is a great collection of essays in different genres. All the narratives chosen focused on personal experiences. The average length of the sample narratives is 216 words. Topics include “My personal experience of helping others,” “A memorable childhood event,” and “My ambition.” The researcher, in collaboration with two experienced English teachers, translated these narratives into English using GT. By analyzing the GT outputs using Shin and Chon’s (2023) error categorization scheme, the three most common types of MT errors, misuse of tenses (N=20), misuse of references (N=8), and mistranslations (N=18), were identified. The following are instances of the three types of GT translation errors.

(1) Misuse of tenses

GT Translation: *“I remember last year when I was still living in Thailand, my parents took my brother and me to visit an orphanage. The children there are very cute, but they are also very pitiful.”*

Corrections: *“I remember last year when I was still living in Thailand, my parents took my brother and me to visit an orphanage. The children there were very cute, but they were also very pitiful.”*

From the above example, it can be seen that GT is unable to maintain consistency in the use of tenses in describing past situations.

(2) Misuse of references

沒想到只是處理玩具太多的問題，最後不但順利解決，同時又能分送給有需要的人，為他們帶來快樂，真是一舉兩得。

GT Translation: *“I didn’t expect that I just solved the problem of*

*too many toys. In the end, not only was **it** successfully solved, but **it** was also distributed to people in need, bringing them happiness. **It** really killed two birds with one stone.”*

Corrections: “*I didn’t expect that, in the end, I not only solved **my problem** but also distributed **the toys** to the people in need, bringing them happiness. **What I did** really killed two birds with one stone.”*

In Example (2), GT uses *it* to replace different nouns in several consecutive sentences, causing confusion and ambiguity. The errors may be associated with the frequent lack of subjects in Chinese sentences. In contrast to English, Chinese is a topic-prominent language. The topic is not a grammatical role but the thing that the sentence is about. The error also reveals the problem that when translating longer texts, GT may not be able to effectively segment the units of meaning and identify the objects it refers to.

(3) Mistranslation

我在週末時把玩具送到孤兒院，送給那些年紀比我小的弟弟妹妹。

GT Translation: “*I took them to the orphanage on the weekend and gave them to **my younger brothers and sisters**.*”

Corrections: “*I took my toys to the orphanage on the weekend and gave them to **the kids**.*”

It can be seen in Example (3) that GT fails to obtain the exact equivalent since “弟弟妹妹” is used figuratively to refer to the kids in the orphanage rather than the children with the same parents as another person.

According to Tian (2004), narratives typically describe events centered on individuals and their actions within a specific time frame. Specifically, narrative essays often utilize third-person pronouns and past tense. In this study, the discovery of the main errors in references

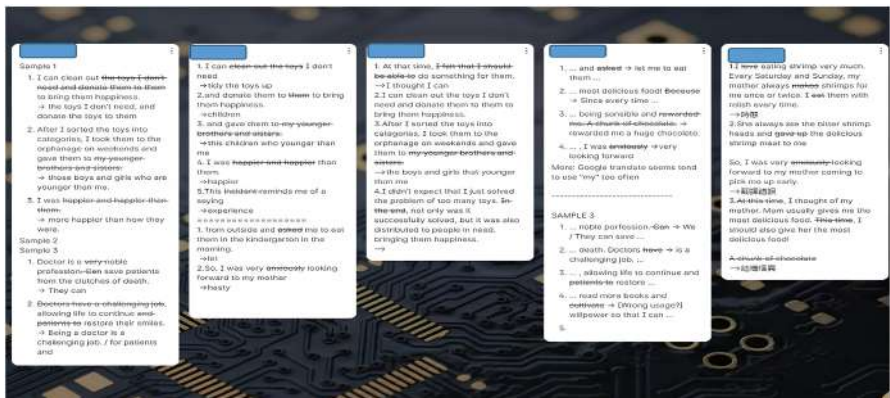
and tenses in MT output reflected the fact that MT may find it difficult to handle narrative discourse.

After identifying the three primary error types in the MT texts, the researcher chose four of the ten collected narratives randomly for the 90-minute training session. This sample size was chosen to ensure sufficient focus and discussion on each text during the limited training period.

At the beginning of MTPE training, the teacher introduced the structure of English narrative. After that, the teacher distributed the handouts with one sample narrative translated by GT to students. Students were invited to detect and post their corrections of the errors in lexicon, syntactic structure, and grammar on *Padlet*, an online board where users can collaborate, reflect, and share ideas (see Figure 1).

Figure 1

Students' Posts of Error Correction on Padlet



After students shared the results of their analysis, the teacher conducted a whole-class discussion and drew their attention to categorized errors and examples from the sample narrative texts. Following that, the teacher distributed three more machine-translated English narrative texts one by one to students for identifying errors within them. The teacher always reviewed the texts with students and

showed them the suggested corrections in the hope that students would become more aware of cross-linguistic differences and the limitations of MT. In addition to illustrating the advantages and disadvantages of MT output during the post-editing training, the teacher also warned the students of the ethical issues that may arise if MT is overused.

Data Analysis

To answer the first research question, we utilized Coh-Metrix, an automated text analysis tool, to examine text features. We carefully selected five indicators of lexical, syntactic, and textual features. Among the lexical features, three indicators were considered: DESWIt, LDTTRc, and PCCNcz. DESWIt represents the average number of letters in words. Lexical diversity, which refers to the variety of words used in a text, is a significant predictor of writing quality. The lexical diversity used in the study was the type/token ratio for content words. This is the ratio of unique words to the total number of words in a text. PCCNcz is an important indicator of text readability component scores. It pertains to the concreteness of words. A higher score indicates a greater percentage of content words that are concrete in meaning rather than abstract.

In terms of syntactic features, SYNMEDwrđ is an indicator of syntactic complexity. A text of high writing quality contains high values of lexical diversity and syntactic diversity. The textual feature CNCCaus is the causal connective incidence. Crossley and McNamara (2009) emphasized the significance of causality in establishing relationships between events and actions. To compare the quality differences between students' L2 texts and revised texts, descriptive statistics and a Wilcoxon signed-rank test were used.

Research question 2 concerns whether students' reliance on MT output changes after MTPE. A content similarity detection system, *Honest Work*, developed by the Center for Taiwan Academic Research Ethics Education, was used to identify similarities between MT output and students' revised English texts in the pretest and posttest. The system identified the most similar sentences in students' texts to those in MT output and then found common word sequences between the

two sentences. By counting the number of words in common word sequences and comparing it with the total number of words in students' texts, how much of students' writing was copied from MT output was determined (Yang & Chou, 2019). Figure 2 shows an example of a comparison result.

Figure 2

A Screenshot Showing the Result of a Text Comparison



A Wilcoxon signed-rank test was used to determine whether there was a significant difference between the pretest and posttest in terms of content similarity.

The third research question examines the relationship between the similarity rate of content and the rate of errors in students' revised texts. To measure the total number of errors occurring in students' texts, QuillBot grammar checker was used. QuillBot is effective in detecting grammar, spelling, and punctuation errors. According to Ho (2022), the QuillBot grammar checker outperformed Grammarly and Ginger, two other popular automated writing tools. The error rates of students' revised texts were obtained by counting the number of errors divided by the total words of the text. Pearson correlation was then used to measure the strength of a linear association between content similarity rate and the rate of errors in the pretest and posttest. Next,

we employed a Wilcoxon signed-rank test to examine if there is any significant difference between the rate of errors in students' L2 texts and revised texts and between the rate of errors in students' revised tests in the pretest and posttest. Finally, all returned questionnaires were analyzed using IBM SPSS 27.

RESULTS

Background Questionnaire Feedback

In order to explore participants' past English writing experience, a background questionnaire was distributed to them in the beginning of the research. Eighty-one percent reflected that they had the experience of writing in English before. More than 70% of students who had English writing experience indicated that they took advantage of MT. Concerning the purposes of using machine translation, students put checking word meaning in the first place (70%), editing self-written English texts (46%) and translating self-written Chinese sentences into English (46%) in the second place, and translating online English texts into Chinese in the third place (20%). With respect to degrees of perceived accuracy of the information provided by MT, the translation of vocabulary ranked highest (80%). It is noteworthy that less than 50% of students agreed that the translated sentences could always/usually show the intended meaning. Similarly, merely 37% of them believed that the grammar in translated texts was usually or always accurate. Overall, while most students held a positive view of the quality of the machine-translated texts, their attitudes became more reserved when confronted with specific issues of concern.

The following section presents the results of each research question. In addition to interpreting statistical data, we apply student feedback from interviews, stimulated recall, and evaluation questionnaires to explain the findings more fully.

Writing Quality

A descriptive analysis and Wilcoxon signed-rank test were

conducted on five indicators between students' English text (L2 text) and revised text (see Table 1).

Table 1

Descriptive Statistics and Wilcoxon Signed-Rank Test

		L2 texts		Revised texts		z	Sig.
		M	SD	M	SD		
Pretest	Word length	3.86	0.81	4.14	0.83	-3.058	<0.001
	Word diversity	0.68	0.52	0.71	0.52	-2.691	0.007
	Word concreteness	-0.51	2.96	-0.71	2.8	0.871	0.414
	Syntactic complexity	0.74	0.30	0.85	0.16	2.8	0.005
	Causal connectives	36	27.7	36	20.87	0.175	0.861
Posttest	Word length	4.16	0.32	4.27	0.36	-2.459	0.014
	Word diversity	0.76	0.06	0.78	0.05	-2.038	0.042
	Word concreteness	-0.07	0.99	-0.25	1.10	-1.395	0.163
	Syntactic complexity	0.78	0.27	0.82	0.22	-1.587	0.112
	Causal connectives	35.35	16.007	34.81	16.19	-0.296	0.767

It can be seen that the revised texts demonstrated a significant increase in both word diversity and sentence length, suggesting that students were influenced by the longer and more complex sentence structures generated by the machine translation system. This indicates that GT may have helped students overcome lexical limitations. This finding aligns with the evaluation questionnaire results, where students strongly agreed that GT texts offered a wider range of vocabulary and were more likely to accept the suggested word choices.

In addition, the background questionnaire revealed high student confidence in the accuracy of MT word translations, suggesting a focus on lexical correspondence between the two language systems during the revision process.

On the other hand, the changes in word concreteness and causal connectives were less marked. This can be attributed to two factors. Firstly, if the original language expressions are relatively abstract, MT may struggle to automatically convert them into more concrete vocabulary. Secondly, narrative texts often contain more descriptive or emotional content, where explicit expressions of causality (such as the use of causal connectives) may not be as essential as they are in explanatory or argumentative texts, thus limiting the potential for improvement in the use of such connectives.

Although the mean scores of word length and word diversity in the revised texts in the post-test were higher than those in the revised texts in the pretest, the difference did not achieve the level of statistical significance, as seen in Table 2. A plausible explanation may be that students' vocabulary size did not increase much during the experiment since the focus of post-editing training is on identifying and correcting grammar errors.

Table 2

Wilcoxon Signed-Rank Test Results of Selected Indicators for Revised Texts in the Pretest and Posttest

Pretest-posttest	z	Sig.
Word length	-0.839	0.399
Word diversity	-1.298	0.193
Word concreteness	-0.159	0.873
Syntactic complexity	-0.558	0.577
Causal connectives	-0.205	0.838

Students’ Reliance on MT Output

The content similarity rate between students’ revised texts and GT texts was examined to show how much students relied on MT. The statistical data was cross-referenced with student responses to enhance the reliability. It can be seen in Table 3 that the mean rate of content similarity in the pretest is over 40%.

Table 3

Content Similarity Rates in the Pretest and Posttest

	Pretest	Posttest
S1	0.00%	9.00%
S2	0.00%	12.00%
S3	0.00%	5.00%
S4	1.00%	9.00%
S5	8.00%	10.00%
S6	8.00%	25.00%
S7	11.00%	94.00%
S8	11.00%	16.00%
S9	26.00%	93.00%
S10	41.00%	63.00%
S11	62.00%	67.00%
S12	77.00%	81.00%
S13	90.00%	92.00%
S14	14.00%	14.00%
S15	15.00%	4.00%
S16	22.00%	0.00%
S17	24.00%	12.00%
S18	25.00%	11.00%
S19	29.00%	8.00%
S20	38.00%	3.00%

	Pretest	Posttest
S21	57.00%	52.00%
S22	58.00%	40.00%
S23	61.00%	23.00%
S24	74.00%	38.00%
S25	75.00%	50.00%
S26	85.00%	22.00%
S27	93.00%	87.00%
S28	97.00%	84.00%
S29	98.00%	17.00%
S30	99.00%	91.00%
	M: 43%	M: 38%
	SD: 34.67%	SD: 33.98%

The reason for heavy reliance on GT texts in the pretest can be attributed to students' inability to manage time effectively and unfamiliarity with the translation process. Some students indicated that they spent too much time on the Chinese version and self-written English texts, leaving insufficient time for final revisions. They chose to improve their revised texts and complete the writing on time by pasting a large portion of the content from GT texts. The Wilcoxon signed-rank test was then used to compare the difference between pretest revised texts and posttest revised texts in terms of content similarity rate. The result shows that there was no significant difference between the two texts ($z = -1.395$, $p = 0.163$). No noticeable decrease or change in the similarity rate in the posttest reflects that students' dependence on GT did not diminish a great deal. Nevertheless, through stimulated recall and interviews, we found that a number of students who kept relying greatly on GT during the post-test carefully examined the GT texts sentence by sentence and considered whether to apply the content into their revised texts. A detailed look at the specific revising moves in screen recording shows that students conscientiously scrutinized and compared sentence structures between their own writing and GT texts. Their decision

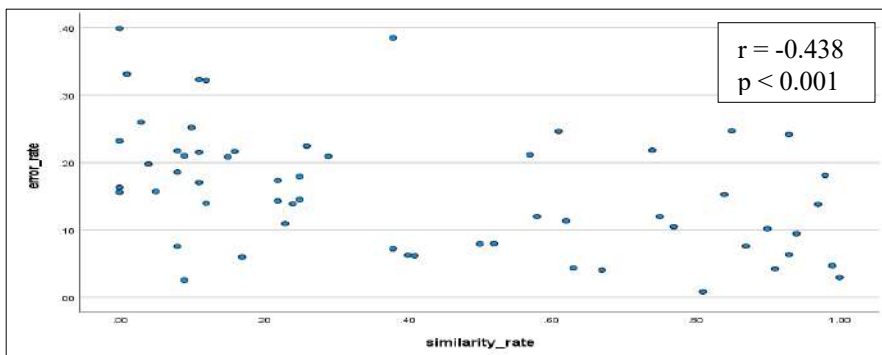
about which strategy to take depended on time constraints, motivational strength, self-belief, and English proficiency. A high percentage of content quoted from GT texts does not necessarily represent student cheating or laziness. As suggested above, some participants whose revised texts had a high content similarity rate (over 60%) reported their trust in the quality of GT texts and approved the revising strategy of using the GT version as a basis, making minor adjustments to align the content more closely with their original intentions. These students actively engaged in the cognitive activities of comparing, judging, and improving. Only a minority of students mentioned quoting a substantial amount of GT content due to low motivation to edit their self-written English texts manually.

The Relationship Between the Content Similarity Rate and the Error Rate

A Pearson correlation coefficient of -0.438 indicates a moderate negative correlation. This negative correlation means that when the content similarity rate becomes higher, the rate of errors becomes lower (see Figure 3). This can explain why some students preferred to quote the content from GT texts. They believed the quality of GT texts is superior to that of their English texts.

Figure 3

Correlation Coefficient Between Content Similarity Rate and Error Rate



On the other hand, the participants whose quotes from the GT texts constituted less than 40% of the whole revised texts in the pretest and posttest expressed skepticism regarding the accuracy rate of GT output because of past experience and the post-editing training in this research. Another group of respondents whose revised texts showed a decrease in similarity with GT texts from pretest to posttest attributed this change to increased confidence in their writing quality in the post-test. Some other participants were concerned that the tone and style of machine-translated articles differed from their own way of expression.

Regarding the rate of errors in both pre- and post-tests, there were no significant differences between L2 texts and revised texts ($z = -1.07$, $p = 0.284$; $z = -0.184$, $p = 0.854$). The result shows that students lacked the ability to correct errors in their revised texts even with reference to GT texts. In the interviews, students reported that while they may have noticed certain expressions looked unusual or odd, they did not know how to correct them. However, it is noteworthy to mention that, as shown in Table 4, the two mean error rates in the posttest were much lower than those in the pretest. Moreover, there is a significant difference between revised texts in pre- and post-tests in terms of error rate ($z = -3.142$, $p = 0.002$).

Table 4

Error Rates in the Pretest and Posttest

	Error rate in pretest L2 texts	Error rate in pretest revised texts	Error rate in posttest L2 texts	Error rate in posttest revised texts
M	20.36%	18.91%	13.09%	12.97%
SD	6.25%	8.85%	9.35%	8.32%

The remarkable decrease in error rates in the revised texts in the post-test suggests that students, after undergoing post-editing training, were more effective in avoiding errors in their writing. The data obtained from the interviews also confirms that students became more attentive to the errors discussed in the post-editing training session, especially the misuse of tenses. In addition, they tended not to use

sophisticated vocabulary or expressions in the Chinese texts to prevent difficulties in translation into English.

Students' Perception of MTPE and Using MT for English Narrative Writing

The first section of the feedback questionnaire deals with students' perceptions of the use of a translation approach in L2 writing. Sixty-four percent of students indicated that generating ideas and writing in Chinese first and then translating the text into English was more efficient than thinking and writing in English ($M=3.72$). In terms of the accuracy of GT-translated segments, 77.7% of students said that GT was more effective in translating vocabulary ($M=3.94$) than translating phrases ($M=3.19$) and sentences ($M=2.92$). On the other hand, less than 30% of students agreed that GT can accurately convey content in L2 ($M=2.94$). As can be seen, the mean scores for these survey items are mostly below 4, indicating that the participants generally did not have a high level of trust in the accuracy of GT texts. Their lower confidence in GT output could be attributed to the type of texts being translated. The content of narratives typically involves more abstract emotions or descriptions of continuous actions and events, making it more challenging for GT to translate precisely. Although students seemed to have some concern about whether the GT texts were reliable, they were positive about the capability of GT in producing the vocabulary (74%, $M=4.08$), phrases (78%, $M=3.97$), and sentence structures (66%, $M=3.72$) they cannot come up with by themselves. With regard to perceived gains, students had moderate attitudes towards gaining the knowledge of English (69%, $M=3.75$) and different ways of structuring texts (60%, $M=3.64$). Concerning their willingness to use GT in the future, only 63% of students expressed intentions to apply it to improve their self-written English texts ($M=3.67$). Students' conservative attitude towards the implementation of this reference tool in writing may be due to the post-editing training received in the workshop. After the error-correction practices, students appear to become much more aware of the errors GT texts may contain.

For the open-ended question concerning students' perceived change in their ways of writing and using GT after MTPE training,

66% of students reported that while they used to rely on GT for searching appropriate vocabulary needed, now they are aware of the possible errors in grammar and misinterpretation. As one interviewee mentioned, *“Before the training, my rough impression of GT was that it might make mistakes, but I didn’t have a clear idea of where the mistakes may occur. After training, I paid attention to the usage of tenses and phrases in the GT text. In addition, I became more alert to whether the intended meaning of my original text was misinterpreted.”* Some participants reflected on the changes in their revision strategies. For example, one interviewee said, *“Before the training, I was not quite sure how to edit my English text. After the training, I paid specific attention to the possible errors introduced in the workshop. In addition, I found that I could reduce the length of units to enhance precision in translation. Sometimes I even refined my Chinese expression in order to obtain more sophisticated English words in GT output.”* Another reported that, *“After the training, I learned that I can improve my English text by comparing it with GT text. Through the comparison, I can carefully make judgments of the proper ways of expression.”* On the other hand, five participants held negative views about the value of post-editing or the quality of GT texts. For instance, according to one interviewee, *“I think whether the content is understandable is more important than how accurate the grammar is.”* Another indicated that *“After the training, I am convinced that GT texts are not quite reliable.”* From the feedback, we can see that students developed a deeper understanding of the capabilities and limitations of GT. The majority of the participants were motivated to critically examine the content of GT texts and modify revision strategies. Nevertheless, some respondents remained doubtful about the value of GT texts and post-editing training.

DISCUSSION

This study aimed to investigate changes in students’ writing quality, dependence on MT, and the rate of grammatical errors before and after receiving post-editing training. In addition to quantitative data analysis, insights into students’ learning experiences and actual

revision processes were gained through screen recording and students' responses to the questionnaire and interview.

With respect to writing quality, we observed that students made more progress in lexical use in comparison with changes in sentence patterns in the revised texts in the pretest and posttest. Word length and word diversity were significantly improved with reference to MT. The finding is consistent with the results reported by Lee (2020), Tsia (2019), and Yang et al. (2023). Students' endorsement of the accuracy of MT, particularly in vocabulary translation, can explain the phenomenon. Intermediate-level students, like those in the present study, frequently express a deficiency in their vocabulary knowledge. They perceive MT as beneficial in suggesting appropriate word choices and enhancing precision in their translations.

The results of the content similarity comparison analysis indicate that there were not significant changes in the similarity rates of texts between the pretest and posttest. While the differences were not great, students exhibited a variety of approaches and considerations when selecting GT content. Regarding the grammatical error rate, there was no significant decrease between students' English texts and revised works in either the pretest or posttest. This indicates that students' ability to detect and correct errors was still limited, which is in line with previous studies highlighting the challenges faced by intermediate-level students in identifying and rectifying errors. (e.g., Lee, 2020). In addition, as suggested in Alrajhi's (2023) study, GT's ability to produce accurate idiomatic expressions and link content ideas in the narrative genre is less satisfying. It's also possible that the sentences in narratives translated by GT contained many semantic or grammatical errors, resulting in a high error rate in intermediate-level students' revised texts. It is worth noting that students' grammatical errors decreased significantly after the post-editing training, which indicates that post-editing was effective in making students aware of possible errors and avoiding them.

Qualitative data revealed that students held mixed feelings about the capabilities of GT, even though they recognized its beneficial impact on the revision process. On the one hand, students affirmed that GT can translate words or sentence structures they might not have

thought of, making the writing process go more smoothly (Cancino & Panes, 2021; Chon et al., 2021; Lee, 2020; Tsai, 2019, 2022). However, they also expressed reservations about the accuracy of GT content. Students realized that their intended meaning sometimes was altered and sentence structures may become quite peculiar when translated from Chinese to English. Moreover, after post-editing training, students had more doubts about the grammatical accuracy of GT content. While some students expressed that their linguistic awareness was raised, others worried that confusion may be caused since they did not have a solid understanding of grammar.

The results of the study suggest the following pedagogical implications. First, teachers should consider how to arrange additional follow-up training or utilize other reference tools, such as corpora or concordancers, to help students internalize new vocabulary and structures observed in GT texts. Previous research has also suggested that covering a single linguistic aspect may require completing multiple writing assignments (Bruton, 2009; Lee, 2020; Sheen, 2007). Providing additional opportunities for students to engage in practice is crucial for fostering a comprehensive understanding of the proper usage of the vocabulary and grammatical structures.

Secondly, for intermediate-level students, teachers should give explicit instruction and organize in-depth lessons based on the categories of MT errors in the two languages being translated. The following teaching suggestions are proposed based on the three error categories identified in the Chinese-English translation focused on in this study. In the unit on the misuse of references, teachers will mark the wrong use of references in the MT output, directing students to read the context and identify the correct antecedents. Through this review, students will become more aware of the importance of coherence and clarity in writing. Furthermore, this topic can be expanded to explore another prevalent challenge faced by intermediate Chinese learners of English—the negative effects of language transfer on the acquisition of English pronouns. In MT output, subject omission may occur because Chinese is a topic-prominent language, whereas English is a subject-prominent language. Teachers can first identify the omitted pronouns in the MT output and highlight their positions, then guide students to observe the Chinese

text and find suitable pronouns or subjects for the corresponding English sentences. This exercise will help students better understand the differences between the two language systems and learn how to revise their L1 text, thereby improving the accuracy of MT. In the unit on the misuse of tenses, teachers will first highlight the verb tense errors and guide students to identify the tenses of other verbs in surrounding sentences. Then, students will be asked to think about and explain why the errors occurred and how they are related to the context. This activity can help students better understand that incongruities in verb tense uses may affect meaning and clarity. Narratives usually describe past events or actions in a sequential manner, so verbs are mostly in the past tense. However, the introduction of the background, descriptions of the influence of past experiences on the present, or the author's current states of mind are often reported in the present tense. Teachers can illustrate these characteristics in narratives to raise students' awareness of how appropriate use of tenses can keep the reader on the hook and make the flow go smoothly. Moreover, it is worth noting that since Chinese is an isolating language, a type of language characterized by having words with little or no inflection, words do not change form to indicate tense. As a result, MT can easily make mistakes when translating between Chinese and English. Other isolating languages such as Vietnamese might also encounter similar issues when translated into English using MT. Teachers can refer to the above teaching suggestions when addressing these challenges.

Lastly, teachers can guide students to explore the reasons why MT misinterprets idioms: (1) some concepts do not exist simultaneously in both language systems, and (2) MT cannot distinguish between literal and figurative meanings. After highlighting the issue, teachers can give students some training in English paraphrasing techniques, such as using synonyms or similar words, changing sentence structures, and providing more detailed descriptions.

CONCLUSION

This study has revealed the benefits of MT use and post-editing

training on helping students improve their writing performance. Students demonstrated more awareness of errors in both their own writing and machine-translated content after the training; they exhibited greater caution when quoting the content from machine-translated texts. Moreover, the study offers some advice for EFL teachers in their grammar and writing courses. These suggestions aim to help students understand that using MT can not only improve the quality of their writing but also provide insights into the importance of text coherence and the positive and negative effects of language transfer.

The significant contributions of this study are as follows:

1. Through both quantitative and qualitative analyses, we not only investigated the differences in students' writing quality before and after using MT and receiving post-editing training but also revealed the internal psychological processes of students during writing. This allows us to identify which text features students prioritized when revising with translation tools and explore the motivations and reasons for focusing on these aspects.
2. In addition, this study provides practical guidelines for educators to integrate MT into writing instruction. It suggests that post-editing training may help balance students' attention to vocabulary and grammatical errors and enhance students' metacognitive awareness of MT's strengths and limitations.
3. The findings of this study regarding content similarity rate counter concerns that MT use might lead to mindless copying. It challenges the simplistic notion that students' dependence is purely for convenience or task completion. The study supports the idea that MT can serve as a tool for enhancing critical thinking skills in language learning.

The current study has several limitations. First, the differences in the prompts for the pretest and posttest may have affected the results of the experiment. Second, there was no control group in the study, and other variables may also have contributed to the outcome. Third,

this was a short-term experimental study focusing on guiding intermediate-level students to understand three common errors that may occur in MT output between Chinese and English. It explored the changes in students' writing performance, writing processes, and psychological attitudes following the stimulation of post-editing training. The research finding may not be generalizable to other studies related to MT involving different language systems.

Future research could investigate the long-term effects of MT and post-editing training on L2 writing by conducting longitudinal studies that track students' language development over an extended period. In addition, exploring how students integrate MT and other AI tools into their writing process, while considering ethical implications, can provide valuable insights for educators and learners. Also, future research should explore task designs that foster autonomous learning and critical thinking while maximizing the potential of AI tools like MT.

REFERENCES

- Ahn, S., & Chung, E.S. (2020). Students' perceptions of the use of online machine translation in L2 writing. *Multimedia-Assisted Language Learning*, 23(2), 10-35.
- Alrajhi, A. S. (2023). Genre effect on Google Translate-assisted L2 writing output quality. *ReCALL*, 35(3), 305-320. <https://doi.org/10.1017/S0958344022000143>
- Bahri, H., & Mahadi, T. (2016). Google Translate as a supplementary tool for learning Malay: A case study at Universiti Sains Malaysia. *Advances in Language and Literary Studies*, 7(3), 161-167. <http://dx.doi.org/10.7575/aiac.all.v.7n.3p.161>
- Bain, A. (1967). *English Composition and Rhetoric* (2nd Edn.). New York: Appleton & Company.
- Baker, C. (2013). *Student and instructor perceptions of online translators in English composition* [Master's thesis, Mississippi State University]. ProQuest Dissertations Publishing.
- Briggs, N. (2018). Neural machine translation tools in the language learning classroom: Students' use, perceptions, and analyses. *JALT CALL Journal*, 14(1), 3-24. <https://doi.org/10.29140/jaltcall.v14n1.221>
- Bruton, A. (2009). Improving accuracy is not the only reason for writing, and even if it were.... *System*, 37(4), 600-613. <https://doi.org/10.1016/j.system.2009.09.005>
- Can, S. (2023). Instructors' perceptions of students' Google Translate use in language learning. *Söylem Journal of Philology*, 474-482. <https://doi.org/10.29110/soylemdergi.1186593>
- Cancino, M. & Panes, J. (2021). The impact of Google Translate on L2 writing quality measures: Evidence from Chilean EFL high school learners. *System*, 98, 102464. <https://doi.org/10.1016/j.system.2021.102464>
- Center for Taiwan Academic Research Ethics Education. (2023). *Honest Work*. <https://ethics.moe.edu.tw/resource/software/>
- Chon, Y. V., Shin, D., Kim, G. E. (2021). Comparing L2 learners' writing against parallel machine translation texts: Raters' assessment, linguistic complexity, and errors. *System*, 96, Article 102408. <https://doi.org/10.1016/j.system.2020.102408>
- Chung, E. S. (2020). The effect of L2 proficiency on post-editing machine translated texts. *The Journal of Asia TEFL*, 17(1), 182-193. <http://dx.doi.org/10.18823/asiatefl.2020.17.1.11.182>
- Chung, E. S., & Ahn, S. (2022). The effect of using machine translation on linguistic features in L2 writing across proficiency levels and text genres. *Computer Assisted Language Learning*, 35(9), 2239-2264. <https://doi.org/10.1080/09588221.2020.1871029>
- Costa, Â., Ling, W., Luis, T., Correia, R., & Coheur, L. (2015). A linguistically motivated taxonomy for machine translation error analysis. *Machine*

- Translation*, 29(2), 127-161. <https://doi.org/10.1007/s10590-015-9169-0>
- Crossley, S. A. & McNamara, D. S. (2009). Computational assessment of lexical differences in L1 and L2 writing. *Journal of Second Language Writing*, 18(2), 119-135. <https://doi.org/10.1016/j.jslw.2009.02.002>
- Eriksson, N. L. (2021). *Google Translate in English-language learning: A study of teachers' beliefs and practices* [Unpublished master's thesis]. Dalarna University.
- Ferris, D., & Roberts, B. (2001). Error Feedback in the L2 Writing Classes: How Explicit Does It Need to Be? *Journal of Second Language Writing*, 10, 161-184. [https://doi.org/10.1016/S1060-3743\(01\)00039-X](https://doi.org/10.1016/S1060-3743(01)00039-X)
- Garcia, I., & Pena, M. (2011). Machine translation-assisted language learning: Writing for beginners. *Computer Assisted Language Learning*, 24(5), 471-487. <https://doi.org/10.1080/09588221.2011.582687>
- Genung, J. F. (1900). *The Working Principles of Rhetoric: Examined in Their Literary Relations and Illustrated With Examples*. Boston: Ginn & Company.
- Godwin-Jones, R. (2019). Telecollaboration as an approach to developing intercultural communication competence. *Language Learning & Technology*, 23(3), 8-28. <http://hdl.handle.net/10125/44691>
- Grenner, E., Åkerlund, V., Asker-Árnason, L., van de Weijer, J., Johansson, V., & Sahlén, B. (2018). Improving narrative writing skills through observational learning: a study of Swedish 5th-grade students. *Educational Review*, 72(6), 691-710. <https://doi.org/10.1080/00131911.2018.1536035>
- Ho, C. C. (2022). The QuillBot grammar checker: friend or foe of ESL student writers? *Journal of Creative Practices in Language Learning and Teaching (CPLT)*, 10 (1), 10-31.
- Jolley, J. R., & Maimone, L. (2015). Free online machine translation: Use and perceptions by Spanish students and instructors. In A. J. Moeller (Ed.), *Learn languages, explore cultures, transform lives* (pp. 181-200). Minneapolis: 2015 Central States Conference on the Teaching of Foreign Languages.
- Kim, E.-Y. (2011) Using translation exercises in the communicative EFL writing classroom. *ELT Journal*, 65(2): 154-160. <https://doi.org/10.1093/elt/ccq039>
- Kliffer, M. (2008). Post-editing machine translation as an FSL exercise. *Porta Linguarum*, 9, 53-67. <https://doi.org/10.30827/Digibug.31745>
- Lee, S.-M. (2020). The impact of using machine translation on EFL students' writing. *Computer Assisted Language Learning*, 33(3), 157-175. <https://doi.org/10.1080/09588221.2018.1553186>
- Lee, S.-M. (2022). Different effects of machine translation on L2 revisions across students' L2 writing proficiency levels. *Language Learning & Technology*, 26(1), 1-21. <https://hdl.handle.net/10125/73490>
- Lee, S. M., & Briggs, N. (2021). Effects of using machine translation to mediate the revision process of Korean university students' academic writing, *ReCALL*, 33(1), 18-33. <https://doi.org/10.1017/S0958344020000191>
- Liu, K., Kwok, H. L., Liu, J., & Cheung, A. K. (2022). Sustainability and influence of machine translation: perceptions and attitudes of translation instructors and

- learners in Hong Kong. *Sustainability*, 14 (11), 6399. <https://doi.org/10.3390/su14116399>
- Moorkens, J. (2018). What to expect from neural machine translation: A practical in-class translation evaluation exercise. *The Interpreter and Translator Trainer*, 12(4), 375–387. <https://doi.org/10.1080/1750399X.2018.1501639>
- Nino, A. (2020). Exploring the use of online machine translation for independent language learning. *Research in Learning and Technology*, 28, 1-32. <https://doi.org/10.25304/rlt.v28.2402>
- O'Brien, S., Simard, M., & Goulet, M.-J. (2018). Machine translation and self-post-editing for academic writing support: Quality explorations. In Moorkens, J., Castilho, S., Gaspari, F. & Doherty, S. (eds.), *Translation quality assessment: From principles to practice*. Cham: Springer, 237–262. https://doi.org/10.1007/978-3-319-91241-7_11
- O'Brien, S. & Ehrensberger-Dow, M. (2020) MT Literacy: A cognitive view. *Translation, Cognition & Behavior*, 3(2), p. 145–164. <https://doi.org/10.1075/tcb.00038.obr>
- O'Neill, E. M. (2016) Measuring the impact of online translation on FL writing scores. *The IALLT Journal*, 46(2), 1–39. <https://doi.org/10.17161/iallt.v46i2.8560>
- Shin, D., & Chon, Y. V. (2023). Second language learners' post-editing strategies for machine translation errors. *Language Learning & Technology*, 27(1), 1–25. <https://hdl.handle.net/10125/73523>
- Sun, D. (2017). Application of post-editing in foreign language teaching: Problems and challenges. *Canadian Social Science*, 13(7), 1-5. <http://dx.doi.org/10.3968/9698>
- Tsai, S.-C. (2019). Using Google Translate in EFL drafts: A preliminary investigation. *Computer Assisted Language Learning*, 32(5–6), 510–526. <https://doi.org/10.1080/09588221.2018.1527361>
- Tsai, S.-C. (2022). Chinese students' perceptions of using Google Translate as a translingual CALL tool in EFL writing. *Computer Assisted Language Writing*, 35, 5-6, 1-23. <https://doi.org/10.1080/09588221.2020.1799412>
- Valijärvi, R. L., & Tarsoly, E. (2019). Language students as critical users of Google Translate: Pitfalls and possibilities. *Practitioner Research in Higher Education Journal*, 12(1), 61–74.
- Weigle, S. C. (2002). *Assessing Writing*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511732997>
- Waller, L., & Papi, M. (2017). Motivation and feedback: How implicit theories of intelligence predict L2 writers' motivation and feedback orientation. *Journal of Second Language Writing*, 35, 54-65. <https://doi.org/10.1016/j.jslw.2017.01.004>
- Zhang, H., & Torres-Hostench, O. (2022). Training in machine translation post-editing for foreign language students. *Language Learning & Technology*, 26(1), 1–17. <http://hdl.handle.net/10125/73466>

ACKNOWLEDGMENT

The author would like to express her sincere gratitude to the anonymous reviewers of the *Taiwan Journal of TESOL* for their insightful and constructive comments, which greatly contributed to the improvement of this paper.

CORRESPONDENCE

Yen-Yu Lin, National Chin-Yi University of Technology, Language Center, Taichung, Taiwan

Email address: yylin@ncut.edu.tw

PUBLISHING RECORD

Manuscript received: July 15, 2024; Revision received: November 27, 2024; Manuscript accepted: January 4, 2025.