



IMAM KHOMEINI
INTERNATIONAL UNIVERSITY



P-ISSN: 2076-5587
Online ISSN: 2076-5905

Google Translate as a Repelling or Propelling Tool: Evidence from Less-Resourced Languages

Zahra Fagher Ajabshir^{ID}

University of Bonab, East Azarbaijan, Iran. fakherzahra@yahoo.com

Article info **Abstract**

Article type: Recently, the use of machine translation (MT) to support second language (L2) writing has increased. Since translation quality via MT matures every year, updated studies are required. The present study explored the quality of MT (*Google Translate*) outputs from Persian to English by comparing them with the texts translated to English by 83 intermediate-level English as a foreign language (EFL) learners. After the instruction on the narrative genre was delivered to the participants, they watched a pictorial narrative prompt and wrote a narrative text directly in English. In the next session, they received training on MT use, watched another narrative prompt, wrote drafts in Persian, and submitted them to MT for translation. A comparison of the texts written directly in English by the participants and MT products showed that the use of MT was beneficial in improving mechanical aspects, lexical sophistication, and some grammatical aspects. Direct L2 writing, on the other hand, rendered better performance in lexical accuracy and some grammatical structures as well as the comprehensibility of the text. Based on the results, pedagogical implications for the use of MT in L2 educational contexts were presented.

Received: 2023/4/30

Accepted: 2024/2/10

Keywords: EFL Learning, Google Translate, L2 Writing, Machine Translation

Cite this article: Fagher Ajabshir, Z. (2024). Google translate as a repelling or propelling tool: Evidence from less-resourced languages. *Journal of Modern Research in English Language Studies*, 11(4), 79-96.

DOI: doi.org/10.30479/jmrels.2023.18709.2205



© The Author(s).

Publisher: Imam Khomeini International University

1. Introduction

With ever-increasing technological advancement, there are currently an array of e-tools widely used in daily life and educational settings. Technologies have offered new affordances in second language (L2) classes for teaching and learning different language skills and components (Mohammad-Salehi & Vaez-Dalili, 2022). Online machine translation (MT) tools have been widely used in L2 settings with a particular focus on writing. MT tools are frequently used by L2 learners due to their convenience of use, fast performance, and cost-effectiveness (Alhaisoni & Alhaysony, 2017; Yoon, 2016). Translation from the first language (L1) to L2 via MT assists learners to retrieve a wide array of words and phrases, thus enabling the production of semantically and syntactically more sophisticated texts in L2 (Tsai, 2019). MT serves as a reference tool for pre-editing, post-editing, and analysis of mistranslations (Chung & Ahn, 2021). It increases students' metacognitive awareness by helping them notice the gaps between their translations and that of MT (Correa, 2014) and monitor and revise their drafts.

While several beneficial effects are associated with the use of MT, there are mixed opinions toward its use with some instructional contexts forbidding or exerting limitations on its use as a primary source of reference. Some instructors believe that MT translations are far below ideal and require careful post-editing to reach an acceptable level (Mundt & Groves, 2016). Moreover, MT translations heavily rely on the subject matter, and the outcomes may not be equally well across different texts. Some pitfalls have been reported such as semantic and syntactic mistranslations (Tsai, 2019), literal translations (Yoon & Chon, 2022), inaccurate cultural and pragmatic understanding, and contextual issues (Tsai, 2019). A further consideration in the use of MT relates to language pairs. It is argued that MT translations between European languages generally yield better accuracy and comprehensibility compared to that in European and Asian languages (Shadiev et al., 2019).

It seems that the ubiquitous nature and omnipresence of MT in L2 classes make its use indispensable. As argued by Ducar and Schocket (2018), the concern is no longer whether instruction should hinder students from consulting such tools, but rather how to assist learners to consider the limitations of these tools and not to move toward MT over-dependency. To this end, ample research is needed to investigate the quality of MT output across varied languages and texts. Moreover, due to increasing improvement in the quality of MT output, more updated studies are required (Lee, 2020). This is especially important for low-resource languages for which there are fewer or limited automatic data resources for machine learning. As stated by Duong (2017), a language is regarded as low-resource for a certain task if it possesses no algorithm using the currently available corpora to do the task at

an acceptable level. Most Asian languages, including Persian, are considered low-resource languages. Although some of these languages possess some digital resources, these resources fail to cover all aspects of the language. This is one of the reasons why MT encounters some struggles in handling Asian languages. According to Mundt and Groves (2015), due to some factors, including the cultural context, different grammar, and different rhetoric styles, MT is less compatible with Asian languages. Given this backdrop, this study adds its contribution to the existing MT research by making a comparative study of GT product and that produced directly by L1-Persian learners English as a foreign language (EFL) students in terms of mechanical, lexical, and grammatical aspects as well as general understandability of the text.

2. Literature Review

Technological online resources have long made a significant contribution to L2 teaching and acquisition, and they have been increasingly accepted and adopted in educational contexts. Currently, the affordances technology contributes to learning in educational contexts is no longer an advantage but rather a necessity (Samani & Noordin, 2020). Since MT has been launched, it has been frequently consulted by L2 students to handle L2 writing challenges, although its use has been occasionally frowned upon by L2 teachers. According to previous research, the pedagogical effectiveness of MT tools is associated with their immediate availability (Zhou et al., 2022), fast translation (Cansino & Panes, 2021), enhanced overall quality of the written output (Stapleton & Kin, 2019), production of more complex and sophisticated texts (Fredholm, 2014), delivering individualized feedback (Lee & Briggs, 2021), and a variety of language pairs involved (Tsai, 2019). The beneficial effects of MT use relate to cognitive and affective aspects of L2 learning as well. It is argued that the use of MT diminishes the cognitive burden and frees up learners' attention from lexico-grammatical micro-level aspects to focus on content and macro-level features of writing, including content and organization (Lee & Briggs, 2021). From the affective perspective, the decreased cognitive load results in a non-threatening learning context, lesser anxiety, and further motivation (Bahri & Mahadi, 2016). On the other hand, the teachers' skepticism toward the use of MT derives from their concerns in terms of lexical and grammatical ambiguity (Somers, 2011), cultural inaccuracies (Mundt & Groves, 2016), and failure to cope with contextual references (Nino, 2008). A further consideration raised by some teachers is plagiarism. There are mixed opinions on whether MT output should be considered the student's authentic work. According to Somers et al. (2006), it is unfair to put equal weight on a text produced by MT with no intellectual effort of the student and a piece of text written originally by the student, allocating time and effort with no MT support.

To shed light on the impact of MT on L2 writing, some previous research has assessed the quality of MT products as compared with students' self-translations or their revisions using MT. In a study by Groves and Mundt (2015), the L1-Malay students were required to submit a text in their L1, and then it was submitted to *Google Translate* (GT), as the most frequently used MT tool. The analysis of the errors in drafts demonstrated that the grammatical level of the GT product was far below the expected level. Serious errors in terms of word choice and sentence structure were found in MT products, suggesting that MT was far from an ideal tool to generate a polished native-like language. However, advantages were found for MT by Tsai (2019) who subjected the Chinese L2 students to writing first in L1, drafting compositions directly in English, and translating the L1 texts into English using GT. The text analysis revealed that the MT texts were characterized by more words, far lesser spelling and grammatical errors, and more sophisticated language choices. It was argued that MT serves as a viable tool to circumvent the challenges involved in conventional processes of language learning, which may yield a huge transformation in L2 learning/teaching practices. Along similar lines, Cansino and Panes (2021) engaged Chilean high-school students in writing short compositions using GT with instruction, GT with no instruction, and control (no access to GT) groups. Both GT groups, while indicating no significant differences, showed performances of better syntactic complexity and accuracy compared to the group with no access to GT. The researchers attributed the lack of difference between the GT groups to the insufficiency of training sessions during which the instruction addressed limited GT features. They recommended that future studies allocate more training sessions.

While previous research has generally documented MT's beneficial effect on L2 writing quality, it is worth noting that the relationship between MT and writing quality is mediated by a number of factors, amongst them language proficiency and text genre. Lee (2020) assigned the low and high-level students to draft L1-Korean texts, translate the scripts into English, and revise the texts using MT. It was found that MT significantly affected low-level learners' correction of lexico-grammatical errors and yielded improved revisions. A similar observation was made by Tsai (2019) and Chon et al. (2022) who advised that MT use can bridge the gap between less and more-knowledgeable students, helping students with limited command of L2 to gain more confidence and motivation. Chung and Ahn (2021) also reported the beneficial impacts of MT for both high and low-level students engaged in writing narrative and argumentative drafts. While narrative MT output contained more lexical diversity, argumentative texts translated by MT included higher syntactic complexity and fluency. Taken together, MT research is rather in its infancy. The conclusions drawn from previous research

are tentative, and ample research is required to draw robust conclusions on the impact of MT (if any) on different aspects of L2 writings. The present study thus aims to assess the quality of MT output and L2 direct writings of students in terms of mechanical, lexical, and grammatical aspects as well as the general understandability of the text. The research questions include the following:

RQ1: Do the L2 narrative texts written directly in English and those originally written in L1 and then submitted to GT for translation show significant differences in terms of mechanical, lexical, and grammatical aspects as well as comprehensibility?

RQ2: What are the students' perceptions toward implementing MT into L2 writing programs?

3. Method

3.1. Participants

The participants were 81 (38 males and 43 females) Iranian college-level EFL students enrolled in the general English course in the autumn semester of 2022. They ranged in age from 19 to 25 ($M = 21.7$ and $SD = 2.3$) with Persian as their first language and an average of 8.3 years of formal English learning. Their baseline English proficiency as assessed by an Oxford Quick Placement Test (OQPT) was intermediate. An ANOVA approved the homogeneity of the participants in terms of their general English proficiency ($M = 31.8$; $SD = 1.93$; $p > 0.05$). From the original pool of 81 students, the drafts of 77 students were analyzed because some students ($n = 4$) failed to attend all the treatment sessions or did not email their self-written or MT output to the teacher. Prior to the treatment, a consent form was filled out by the participants, and they were assured of anonymity.

3.2. Materials and Procedure

The machine translation tool used in this study was Google Translate. It is a user-friendly and convenient technological tool commonly used via web browsers or Android interfaces for translating into over 100 languages. In addition to text translation, it offers a number of possibilities, including language detection, offline translation, spelling check, as well as handwritten and multimedia translation.

The treatment followed four phases: (1) instruction on the narrative genre, (2) direct writing (DW) in English, (3) GT training, (4) writing in Persian and uploading it for GT translation, and (5) a survey questionnaire. In phase 1, the instruction on the narrative genre was delivered by the teacher. The teacher asked volunteer students to narrate a true story from their own experience, followed by some questions about the stories and the discussion of the structure of the story and the different features shaping it. The instruction

was mainly centered on discussing various elements in narratives, including setting, characters, relations, plot (sequence of events), moves, etc.

In phase 2, the students were assigned a narrative writing task. According to Cancino and Panes (2021), narrative writing helps students explore different characters and settings and express their stories in an organized way. Once a narrative task is completed in response to a visual stimulus, it can produce useful outcomes. A narrative prompt in the form of sequenced pictures was displayed on a projector, and the participants were asked to develop a 200-300-word story based on the pictures shown. The prompt depicted a crow that has a piece of cheese in its mouth and is sitting on a tree branch under which there is a fox, and the fox is trying to snatch the cheese from the crow (See Appendix 1). According to Chung and Ahn (2021), the failure to control the use of external resources during task completion may undermine the validity of the claim that the improvement in the final product (if any) is the effect of MT. Accordingly, during the task completion, which took 30 minutes, the participants were not allowed to consult any online/print dictionaries or any web-based resources. They wrote the text on their laptops or cellphones, and upon the completion of the task emailed the drafts to the teacher (researcher).

In phase 3, they attended a GT training session. The teacher elaborated on the GT, its different affordances, advantages, and limitations. Using the video projector, a sample text was uploaded to GT, followed by a detailed discussion of various aspects of the GT translated version of the text and different features of GT, including translation history, definitions of single words, back translation, etc. The students were also encouraged to translate sample texts and check different GT aspects as well as mistranslations. The teacher walked around the class and acted as a facilitator, providing support if needed.

In step 4, a further narrative prompt was offered (See Appendix 1). It depicted a boy who finds an injured bird while crossing the road and takes it home and protects it. After a while, when he opens the window and sees birds flying freely, he decides to release the bird. The sequenced pictures were shown on the projector. Both narrative prompts in phases 2 and 4 were adopted from www.pinterset.com, a website that provides free sequence writing prompts for pre-intermediate and intermediate-level students. The students were required to write a 200-300-word narrative text in Persian and then upload it to GT for translation. It took 30 minutes to complete the task. They saved their drafts written in Persian as well as GT translations and emailed them to the teacher-researcher for analysis.

Finally, in step 5, a questionnaire survey including nine multiple-choice items as well as an open-ended question was administered. The multiple-choice items mainly elicited the participants' responses on the

language pairs they use in MT, different aspects of L2 writing that they used MT or felt it would be helpful, and their willingness to keep using it in the future. The open-ended question required them to comment on the areas not covered in multiple-choice items. The questionnaire items were adapted from earlier studies (e.g., Clifford et al., 2013; Nino, 2008; Tsai, 2019). To ensure content validity, the questionnaire was checked by two experienced EFL teachers, and some items were revised. The Cronbach's alpha reliability of the questionnaire (0.79) was found to be acceptable.

3.3. Text Assessment

The teacher-as-researcher and an EFL rater performed the analysis and scoring of the DW and GT texts. Both raters hold Ph.D. in Teaching English as a Foreign Language (TEFL) and had 12 and 15 years of teaching English, respectively. Prior to the assessment, the raters were briefed on the assessment process and received a detailed explanation of the rubric. They assessed sample texts and discussed the ambiguities and areas of uncertainty. The inter-rater reliability was acceptable as confirmed by an alpha coefficient of .81.

Adapted from previous studies (e.g., Briggs, 2018; Lee, 2020; Stapleton & Kin, 2019), four writing quality measures were used for the analysis of the drafts: (a) spelling and punctuation, (b), vocabulary, (c) grammar, and (d) comprehensibility. The units of analysis differed for each measure. Word served as the unit of analysis of spelling and vocabulary, sentences for grammar, and paragraphs for comprehensibility. Each measure was assessed by a 4-point Likert scale ranging from "poor" to "excellent". To convert the ordinal data in the Likert scale into measurable scales, the corresponding values were assigned for each response type, that is 4 for "excellent", 3 for "good", 2 for "fair" and 1 for "poor" (See Appendix 2 for the descriptions of each level).

Content and organization, as argued by Qin and Uccelli (2016), are important predictors of written quality. However, because these two aspects lie out of the capabilities of MT and deviate from this study's focus, they were not assessed.

4. Results and Discussion

4.1.1. Writing Assessment across Two Translation Modes

Table 1 shows a series of independent t-tests for writing measures of DW and GT versions. The parameters under investigation for text quality yielded different scores across two translation modes. Mechanical aspects, including spelling and punctuation, received better mean scores by MT than DW (M = 4.63 for DW; M = 4.79 for GT for spelling; M = 4.17 for DW; M=4.95 for GT for punctuation). The t-test verified that MT yielded better performance lexical items in spelling ($t = 3.247$, $p < 0.05$) and punctuation ($t = 5.061$, $p < 0.05$). Drawing on Cohen's (1998) criterion, the values found for the effect size (Partial $\eta^2 = 0.89, 0.91$) were high, which means that the

magnitude of the difference between the two translation modes was high concerning mechanical aspects.

Table 1

Independent T-Test for Writing Measures of Direct Writing (DW) and Google Translated (GT) Versions

		DW texts		GT texts		t	Sig.	Partial η^2
		M	SD	M	SD			
Mechanics	Spelling	4.63	1.26	4.79	1.73	3.247	0.000	0.89
	Punctuation	4.17	1.06	4.95	1.49	5.061	0.000	0.91
Lexicon	Lexical accuracy	3.85	1.01	3.78	0.91	2.431	0.001	0.68
	Lexical sophistication	3.71	1.73	3.79	1.04	2.009	0.003	0.81
Grammar	Grammatical accuracy	3.74	1.46	3.79	1.12	1.378	0.17	0.42
Comprehensibility	Idiomatic expressions	3.69	1.13	3.51	0.93	3.484	0.000	0.53
	Contextual understanding	3.78	1.21	3.71	1.41	1.041	0.000	0.81

With regard to lexical measures, GT obtained higher mean scores than DW in vocabulary sophistication ($M = 3.71$ for DW; $M = 3.79$ for GT), but not in accuracy ($M = 3.85$ for DW; $M = 3.78$ for GT). The t-test results verified a significant difference between GT and DW scores in vocabulary accuracy ($t = 2.431$, $p < 0.05$) with a moderate effect size (Partial $\eta^2 = 0.68$) as well as sophistication ($t = 2.009$, $p < 0.05$) with a large effect size (Partial $\eta^2 = 0.81$).

Concerning grammatical accuracy, no significant difference appeared between GT and DW modes ($t = 1.378$, $p > 0.05$). The qualitative analysis of the drafts, however, suggested some differences in translations of some grammatical aspects across translation modes. GT was found to retrieve correct translations for propositions, plural s, definite/indefinite articles, and coordinated structures while DW yielded correct embedded structures.

Regarding comprehensibility, as measured by idiomatic expressions and contextual understanding, the results showed better performance for DW. The mean scores of idiomatic expressions for DW and GT were 3.69 and 3.51, respectively. Similarly, DW and GT obtained 3.78 and 3.71 mean scores for contextual understanding, respectively. The text results revealed the outperformance of DW compared to GT in idiomatic expressions ($t = 3.484$, $p < 0.05$) with a moderate effect size (Partial $\eta^2 = 0.53$) and contextual understanding ($t = 1.041$, $p < 0.05$) with a large effect size (Partial $\eta^2 = 0.81$).

4.1.2. Analysis of Questionnaire Responses

Adopting a mixed-method design, this study surveyed the students' perceptions of MT. A Persian version of an anonymous survey questionnaire was distributed among the students to elicit their attitudes toward MT. Table 2 shows the responses to questionnaire items. The first item required the respondents to report their previous experience of using MT (if any).

Interestingly, 87% of the respondents reported that they were familiar with MT and had previously used it.

The second item required the respondents to specify what languages they have ever used MT for. English received the highest scores (89%), followed by French (5%), Spanish (3%), Chinese (1%), and other languages (2%). This is not surprising, as English features prominently in Iran's official educational system and serves as an obligatory course that should be taken by all undergraduate students.

Table 2

Response to Questionnaire Survey

1. Have you ever used MT to support the language learning process prior to this course?				
Yes		No		
87%		13%		
2. For which languages you have used MT?				
English	French	Chinese	Spanish	Other
89%	5%	1%	3%	2%
3. Do you find MT in your English learning... helpful?				
Always	Often	Sometimes	Rarely	Never
32%	21%	26%	16%	5%
4. How accurate is the MT that you use for translation into English?				
Always	Often	Sometimes	Rarely	Never
10%	18%	52%	13%	7%
5. For what levels of writing do you use MT?				
Individual words	Short phrase	Sentences	Paragraphs	Longer texts
23%	19%	18%	21%	19%
6. For what stages of writing, do you use MT?				
Pre-writing	Writing	Editing	Revising	
11%	23%	21%	38%	7%
7. In which aspects of writing, do you find MT helpful?				
Punctuation and spelling	Vocabulary	Grammar	Idiomatic expressions	Content organization and
11%	57%	20%	4%	6%
8. In which of the following ways you find MT helpful?				
Building confidence	Production of a more native-like language	Saving time	Improvement of your grades	Saving cost
23%	19%	20%	18%	20%
9. Do you continue using MT?				
Always	Often	Sometimes	Rarely	Never
21%	23%	32%	15%	9%

10. Please mention other ways you have found MT helpful/harmful not listed above.

To the question that if MT was helpful in students' English learning, 32% answered *always*, 21% chose *often*, 26% *sometimes*, 16% rarely, and 5% *never*. This suggests that most of the students perceived MT as helpful in learning English.

Concerning the accuracy of English translations by MT, more than half of the students reported that it was *sometimes* accurate (52%), 10% responded *always*, 18% *often*, 13% *rarely*, and 7% *never*. It seems likely that the answers to this question should be regarded as tentative. The experimentation in this study was limited to narratives which, compared to other genres, is regarded as a rather simple genre that may pose fewer challenges to MT.

The percentages reported for the responses in item 5 were roughly similar. Twenty-three percent of the students used MT for the translation of individual lexical items, 19% for short phrases, 18% for sentences, 21% for paragraphs, and 19% for longer texts. This suggests that the participants used MT almost equally for the translation of different building blocks of text at different levels.

Regarding item 6, 21% of the respondents asserted that they used MT for editing, 23% for writing, 38% for revising, and 11% for pre-writing. Most of the students thus did not use MT for writing from scratch but rather for the comparison of their drafts and that of MT to make revisions and produce high-quality texts.

When surveyed to identify the specific ways the respondents deemed MT useful (Q7), 57% reported that MT helps increase vocabulary, followed by grammar (20%), punctuation and spelling (11), content and organization (6%), and idiomatic expressions (4%).

For item 8, the respondents assigned roughly similar weights to the responses. They hold different attitudes toward the ways that MT proves beneficial. For 23% MT helped build confidence; for 19% it was helpful in the production of a more native-like language; 18% responded that it improved their grades; for 20%, it saved time; and 20% reported that it was cost-effective.

The item that asked respondents whether they wish to keep using MT in the future received different responses. Thirty-two percent of the participants responded that they wish to use MT *sometimes*, 23% *often*, 21% *always*, 15% *rarely*, and 9% *never*.

As for the open-ended questions, 53 students wrote comments and suggestions. The pattern of responses showed that for 21%, MT's fast performance helps the smooth flow of the writing task. For 17%, it was especially useful for problematic phrases and clauses, and 25% argued that it raised their consciousness of certain grammatical structures.

Meanwhile, 16% and 14% argued that MT production was far from ideal, and there is still room for amendments in grammar and idiomatic expressions. Seven percent of the respondents hold negative attitudes toward MT and believed that students must complete tasks with no reliance on MT.

4.2. Discussion

This study explored how the use of MT as compared with DW affects the mechanical, lexico-grammatical, and comprehensibility of narrative texts written by intermediate-level EFL learners. Overall, the use of MT was beneficial in measures related to mechanical aspects, lexical sophistication, and some grammatical aspects (e.g., coordination, plural s, and definite/indefinite articles), lending support to some earlier studies that have documented the beneficial outcomes associated with the use of MT (e.g., Chung & Ahn, 2021; Lee, 2020; Petrucci et al., 2018; Stapleton & Kin, 2019; Tsai, 2019). Direct L2 writing, on the other hand, rendered better performance in lexical accuracy and some grammatical structures (e.g., subordination) as well as comprehensibility of the text.

Not surprisingly, for mechanics, MT yielded better outcomes. As argued by Groves and Mundt (2015), due to the dichotomous nature of mechanical aspects, MT manipulates spelling errors very carefully. Once encountered with random non-word or real-word spelling errors, it offers several spelling alternatives (Dong et al., 2019). It also handles typological errors such as duplication, addition, omission, substitution, etc. (Chakravarthi et al., 2021). Due to error-recovery potential, MT corrects spelling mistakes in the input before performing the translation. Contrary to the earlier versions of translation software that were prone to some spelling errors derived from L1 misspelling, the increasing maturity of the software enabled retrieving the correct translation of words written incorrectly in the source language.

While direct L2 writings were generally marked with more accurate choices of lexical items, GT translations were characterized by a more sophisticated lexicon. Further inspection of the written drafts revealed that the DW drafts included the basic, common, and readily available lexical items. Rather than using a variety of lexical items to enrich their texts, the students tended to rely on their “lexical teddy bears” (Hasselgren, 2007). This is consistent with the findings of Uzawa (1996), Stapleton and Kin (2019), and Lee (2021) who found lesser diversity and sophistication in directly translated texts. In contrast, GT translations were characterized by more diverse, lexically rich, and advanced words. It seems that L1 served as a mediator, and the source texts submitted to GT assisted in retrieving a wide range of vocabulary items. According to Stapleton and Kin (2019), in some cases, writing in L1 followed by GT translation not only yields accurate translation but also one with further sophistication and refinement.

Regarding grammar, while GT failed to successfully handle some syntactic structures (e.g., embedding), it could efficiently deal with some structures that L2 students often struggle with (e.g., prepositions, article use, verb tense, and agreement). An area of challenge for GT was related to gendered third-person pronouns. Persian is a gender-neutral language, using a single pronoun (او) for both masculine and feminine genders. Contrary to Persian, the English language possesses gender-specific pronouns, which caused GT not to make correct pronoun distinctions, yielding incorrect gendered pronouns in some cases. It is argued that GT fails to make subtle distinctions between L1 and L2 forms, suggesting that the software is not capable of parsing some syntactic structures, which results in ambiguity (Petrucci et al., 2018). While some empirical studies (e.g., Garcia & Pena, 2011; O'Neill, 2016; Tsai, 2019) provided supportive evidence for the efficacy of GT in retrieving correct syntactic structures, some other studies (e.g., Chon et al., 2021; Groves & Mundt, 2015;) argued that MT failed to retrieve correct grammatical forms for some structures. A possible reason for grammatical mistranslation, according to Chon et al. (2021) may relate to students' tendency to write in formats and structures that do not correspond to the features of the target language.

While DW texts were generally characterized by simple, more familiar, and basic lexicon and syntax, they featured more clarity. This is not surprising as MT offers the words' literal translations and does not efficiently handle words' connotations, denotation, and context (Chung & Ahn, 2021). Moreover, GT failed to provide equally well idiomatic expressions, which undermined the comprehensibility of the texts. GT texts contained imprecise literal translations for some idioms, suggesting that GT generally provides the general, conventional, and literal meaning of the word rather than its figurative meaning (Ducar & Schocket, 2018). According to Stapleton and Kin (2019), GT rules are based on word frequencies and associations and prioritize probability over contextual metaphorical meaning. This may render mistranslations that are generically and contextually inappropriate in the target context, leading to ambiguity and a lack of comprehensibility.

The participants' responses to the survey questionnaire suggest that they generally hold positive attitudes toward MT. This corroborates the findings of Nino (2008), Clifford et al. (2013), Tsai (2019), and Murtisari et al. (2019), suggesting that the students while acknowledging the deficiencies of MT, perceived it as pedagogically effective in terms of improved overall writing quality, better lexical choices, and structural enhancement. According to Murtisari et al. (2019), students used MT as a replacement for paper dictionaries for the translation of words and short phrases. Those who use MT for longer stretches of text may fall short of developing a good command of the L2 writing skill. Overreliance on MT may encourage avoidance of L2

cognitive processes, which is fundamental to L2 acquisition. In a similar vein, Lee (2022) argued that the students in his survey, perceived MT as useful from L2 to L1 for longer texts to gain a quick grasp of the text. They rarely considered L1 to L2 longer texts' translation by MT helpful. Students thus are informed about the instructional advantages of MT. At the same time, they are aware of the shortcomings of this technological platform. This, as stated by Clifford et al. (2013) poses the need for a critical assessment of MT by students, which remains a skill to be fostered as part of the educational curriculum.

5. Conclusions and Implications

The present study found evidence for the effectiveness of GT in some aspects of writing quality. Findings revealed that GT is beneficial in generating a writing product with improved mechanical aspects, enriched vocabulary, and grammatical aspects. This contributes to the evidence suggesting the efficacy of GT as an educational platform that fosters enhanced text quality. Considering the findings of this study, EFL teachers and educationalists are recommended to integrate GT as an essential component of learning. In addition to contributing to students' digital literacy (Ducar & Schocket, 2018), MT is especially helpful for beginner and intermediate-level students whose requirements are basic which can be easily handled by GT. Although MT in general and GT, in particular, are at present far from serving the role of a fluent human translator, these technological tools by no means serve as a vital referencing tool for students (Chon et al., 2022).

Given the educational effectiveness of MT tools, it is suggested to use these platforms as additional language learning tools to support improved writing. As argued by Lee (2020), MT should not be considered a replacement for conventional classroom practices nor should language learners be inhibited from using its potential. It should be used in a way not to undermine actual learning processes. A growing concern among teachers is that once correct translations are readily available by uploading texts in L1, the learners' motivation to learn to write in L2 diminishes (Stapleton & Kin, 2019). Such dependency on MT may lead to shallow rather than deep literacy (Groves & Mundt, 2015). Considering the ubiquitous use of MT and the instructors' concerns, teachers are advised to expand students' knowledge of the limitations of this tool and codes of practice and use it with caution. Students can use it as a post-writing editing tool and/or for looking up single words or phrases rather than long stretches of text (O'Neill, 2019). Learners should realize that capturing subtleties across languages is a challenging task, and MT might fail to appropriately convey the cultural concepts, beliefs, and values, of the target text. Strategy-based instruction is recommended to be delivered to students to get an awareness of how to resolve the problems associated with the mistranslations of MT.

Although this study verified the enhanced quality of some writing parameters following the use of MT, the findings are tentative as they address a narrow group of students, a single genre, and a single language being translated. Undoubtedly, there are areas of MT research yet to be explored. Future longitudinal studies incorporating varied MT tools on students of different proficiency levels in less-explored linguistic contexts will offer valuable insights into how MT can be effectively incorporated in L2 classes.

Acknowledgements

We would like to thank the participants who agreed to take part in the present study.

References

- Alhaisoni, E., & Alhaysony, M. (2017). An investigation of Saudi EFL university students' attitudes towards the use of Google translate. *International Journal of English Language Education*, 5(1), 72-82. <https://doi.org/10.5296/ijele.v5i1.10696>.
- 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. <https://doi.org/10.7575/aiac.all.v.7n.3p.161>.
- Briggs, N. (2018). Neural Machine Translation Tools in the Language Learning Classroom: Students' Use, Perceptions, and Analyses. *JALT Call Journal*, 14(1), 2-24. <https://doi.org/10.29140/jaltcall.v14n1.221>.
- 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>.
- Chakravarthi, B. R., Rani, P., Arcan, M., & McCrae, J. P. (2021). A survey of orthographic information in machine translation. *SN Computer Science*, 2(4), 330. <https://doi.org/10.1007/s42979-021-00723-4>
- Chon, Y. V., Shin, D., & Kim, G. E. (2021). Comparing L2 learners' writing against parallel machine-translated texts: Raters' assessment, linguistic complexity and errors. *System*, 96, 102408. <https://doi.org/10.1016/j.system.2020.102408>.
- 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>.
- Clifford, J., Merschel, L., & Munné, J. (2013). Surveying the landscape: What is the role of machine translation in language learning? *Revista D'innovació Educativa*, 10, 108-121.
- Correa, M. (2014). Leaving the "peer" out of peer-editing: online translators as a pedagogical tool in the Spanish as a second language classroom. *Latin*

- American Journal of Content & Language Integrated Learning (LACLIL)*, 7, 1–20. <https://doi.org/10.5294/laclil.2014.7.1.1>.
- Dong, R., Yang, Y., & Jiang, T. (2019). Spelling correction of non-word errors in Uyghur–Chinese machine translation. *Information*, 10(6):202. <https://doi.org/10.3390/info10060202>.
- Duong, L. (2017). Natural language processing for resource-poor languages. *University of Melbourne*.
- Ducar, C., & Schocket, D. (2018). Machine translation and the L2 classroom: Pedagogical solutions for making peace with Google translate. *Foreign Language Annals*, 51(4), 779–795. <https://doi.org/10.1111/flan.12366>.
- Fredholm, K. (2019). Effects of Google Translate on lexical diversity: Vocabulary development among learners of Spanish as a foreign language. *Revista Nebrija de Lingüística Aplicada a la Enseñanza de Las Lenguas*, 13(26), 98–117. <https://doi.org/10.26378/rnlael1326300>.
- Garcia, I., & Pena, M. I. (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>.
- Hasselgren, A. (2007). Lexical teddy bears and advanced learners: A study into the ways Norwegian students cope with English vocabulary. *International Journal of Applied Linguistics*, 4, 237–258. <https://doi.org/10.1111/j.1473-4192.1994.tb00065.x>.
- Lee, S. (2020). The impact of using machine translation on EFL students' writing. *Computer Assisted Language Learning*, 33(3), 157–75. <https://doi.org/10.1080/09588221.2018.1553186>.
- Lee, S. (2022). An investigation of machine translation output quality and the influencing factors of source texts. *ReCALL*, 34(1), 81–94. <https://doi.org/10.1017/S0958344021000124>.
- Lee, S., & 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>.
- Mohammad-Salehi, B., & Vaez-Dalili, M. (2022). Examining EFL teachers' perceptions of technological pedagogical content knowledge and Web 2.0 technologies using a structural equation modeling technique. *Journal of Modern Research in English Language Studies*, 9(2), 51–76. <https://doi.org/10.30479/jmrels.2021.14550.1779>.
- Mundt, K., & Groves, M. (2016). A double-edged sword: the merits and the policy implications of Google Translate in higher education. *European Journal of Higher Education*, 6(4), 387–401. <https://doi.org/10.1080/21568235.2016.1172248>.
- Murtisari, E., Widiningrum, R., Branata, J., & Susanto, R. (2019). Google translate in language learning: Indonesian EFL students' attitudes. *The Journal of Asia TEFL*, 16, 978–986.

- <https://doi.org/10.18823/asiatefl.2019.16.3.14.978>.
- Nino, A. (2008). Evaluating the use of machine translation post-editing in the foreign language class. *Computer Assisted Language Learning*, 21(1), 29-49. <https://doi.org/10.1080/09588220701865482>.
- O'Neill, E. (2019). Training students to use online translators and dictionaries: the impact on second language writing scores. *International Journal of Research Studies in Language Learning*, 8, 47-65. <https://doi.org/10.5861/ijrsl.2019.4002>.
- Petrucci, G., Rospocher, M., & Ghidini, C. (2018). Expressive ontology learning as neural machine translation. *Journal of Web Semantics*, 52-53, 66-82. <https://doi.org/10.1016/j.websem.2018.10.002>.
- Qin, W., & Uccelli, P. (2016). Same language, different functions: A cross-genre analysis of Chinese EFL learners' writing performance. *Journal of Second Language Writing*, 33, 3-17. <https://doi.org/10.1016/j.jslw.2016.06.001>.
- Samani, E., & Noordin, N. (2020). Getting connected with Facebook Messenger: Exploring meaningful interactions through online chats in the ESL context. *Journal of Modern Research in English Language Studies*, 7(3), 23-44. <https://doi.org/10.30479/jmrels.2020.11985.1492>.
- Shadiev, R., Sun, A., & Huang, Y. M. (2018). A study of the facilitation of cross-cultural understanding and intercultural sensitivity using speech-enabled language translation technology. *British Journal of Educational Technology*, 50(3), 1415-1433. <https://doi.org/10.1111/bjet.12648>.
- Somers, H. (2011). Machine translation: History, development, and limitations. In K. Malmkjær, & K. Windle (Eds.), *The Oxford Handbook of translation studies* (pp. 427- 440). Oxford University Press.
- Somers, H., Gaspari, F., & Niño, A. (2006). Detecting inappropriate use of free online machine-translation by language students: A special case of plagiarism detection. In *Proceedings of the 11th Annual Conference of the European Association for Machine Translation* (pp. 41-48), Norway.
- Stapleton, P., & Kin, B. (2019). Assessing the accuracy and teachers' impressions of Google Translate: A study of primary L2 writers in Hong Kong. *English for Specific Purposes*, 56, 18-34. <https://doi.org/10.1016/j.esp.2019.07.001>.
- 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>.
- Uzawa, K. (1996). Second language learners/ processes of L1 writing, L2 writing, and translation from L1 into L2. *Journal of Second Language Writing*, 5, 271-294. [https://doi.org/10.1016/S1060-3743\(96\)90005-3](https://doi.org/10.1016/S1060-3743(96)90005-3).

Appendix 2

Rubrics for the Assessment of Written Texts

Mechanics	Spelling/ Punctuation	<p><i>Excellent:</i> Indicates mastery of conventions; no errors of spelling, punctuation, capitalization</p> <p><i>Good:</i> Indicates an acceptable knowledge of conventions; less errors of spelling, punctuation, capitalization</p> <p><i>Fair:</i> Indicates less knowledge of conventions; some errors of spelling, punctuation, capitalization</p> <p><i>Poor:</i> Indicates poor knowledge of conventions; much errors of spelling, punctuation, capitalization</p>
	Lexical accuracy	<p><i>Excellent:</i> Uses a wide range of vocabulary appropriately and correctly</p> <p><i>Good:</i> Uses a roughly acceptable range of vocabulary most often appropriately and correctly with a few errors</p> <p><i>Fair:</i> Uses a limited range of words with many minor errors and a few major errors</p> <p><i>Poor:</i> Uses a very limited range of words with lots of minor and major errors</p>
Lexicon	Lexical sophistication	<p><i>Excellent:</i> Uses a wide range of less-frequent and advanced lexicon</p> <p><i>Good:</i> Uses a roughly acceptable range of less-frequent and advanced lexicon</p> <p><i>Fair:</i> Uses a limited range of less-frequent and advanced lexicon</p> <p><i>Poor:</i> Uses a very limited range of less-frequent and advanced lexicon</p>
	Grammatical accuracy	<p><i>Excellent:</i> Error-free or minimal number of errors (including common difficulties such as tense, subject-verb agreement, articles, plurals, complex sentences, etc.)</p> <p><i>Good:</i> Some error-free sentences but only minor errors in most sentences</p> <p><i>Fair:</i> Most sentences include minor errors with some major errors</p> <p><i>Poor:</i> Many errors throughout both minor and major</p>
Comprehensibility	Idiomatic expressions	<p><i>Excellent:</i> All the idioms used were accurate and meaningful.</p> <p><i>Good:</i> Some of the idioms used were accurate and meaningful.</p> <p><i>Fair:</i> The idioms used were seldom accurate and meaningful.</p> <p><i>Poor:</i> The idioms used were never accurate and meaningful.</p>
	Contextual understanding	<p><i>Excellent:</i> Completely clear and understandable considering the given context</p> <p><i>Good:</i> Mostly understandable with some minor ambiguities</p> <p><i>Fair:</i> Partially understandable with a few major ambiguities or incomprehensible expressions</p> <p><i>Poor:</i> Some sentences are understandable, but much of the script is beyond comprehension</p>