



Castledown

 OPEN ACCESS

# Australian Journal of Applied Linguistics

ISSN 2209-0959

<https://www.castledown.com/journals/ajal/>

*Australian Journal of Applied Linguistics*, 5(3), 135–157 (2022)

<https://doi.org/10.29140/ajal.v5n3.53si3>

## Online Machine Translation for L2 Writing Across Languages and Proficiency Levels



ANTONIE ALM<sup>a</sup>

YUKI WATANABE<sup>a</sup>

<sup>a</sup>*University of Otago, New Zealand*

[antonie.alm@otago.ac.nz](mailto:antonie.alm@otago.ac.nz); [yuki.watanabe@otago.ac.nz](mailto:yuki.watanabe@otago.ac.nz)

### Abstract

Using machine translation (MT) tools for language learning has become a common practice among language students in recent years. Studies have investigated how students use MT, how students and teachers perceive its benefits and drawbacks and how helpful it is for language learning. These studies indicate that students think MT tools are helpful in L2 writing due to their quick and easy access and use them in many aspects of L2 writing, such as vocabulary search, grammar checking, and writing revisions. However, concerns for the accuracy of outputs, the effectiveness of MT for language learning and academic integrity are shared among students and teachers. This present study is based on a survey of 12 teachers and 150 students across five different languages and three proficiency levels at a tertiary institution in New Zealand. The quantitative and qualitative data were analysed to compare MT use and perceptions among proficiency levels and languages as well as between teachers and students. The findings reveal patterns that indicate different practices and perceptions between students of non-alphabet-based and alphabet-based languages. The analysis also demonstrates correlations between advanced-level students and more sophisticated and critical use of MT. With the proliferation of MT tools with neural machine translation (NMT) systems and improvement in their accuracy, the findings of this study contribute to the understanding of MT literacy to ensure effective and critical use of MT with an awareness of unique contexts and expectations of L2 writing in different language courses.

**Keywords:** machine translation, student experiences, teacher experiences, proficiency levels, language pairs

**Copyright:** © 2022 Antonie Alm, Yuki Watanabe. This is an open access article distributed under the terms of the Creative Commons Attribution Non-Commercial 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within this paper.

## Introduction

Douglas Adam's science fiction classic *The Hitchhiker's Guide to the Galaxy* (1979) predicted several technological advances. One of them is the Babel Fish. Described as "the oddest thing in the Universe," the small, yellow and leech-like creature hid in people's ears, giving them the power to "instantly understand anything said to [them] in any form of language" (p. 51). Common to many science fiction works, the idea of a universal translator has long fascinated humankind. However, when language students discovered two decades later an online translator of the same name, language teachers reacted with dismay and were quick to point out the limitations of machine translation engines. Luton (2003), for example, commented, "they cannot recognize idioms and tend to translate them word for word, producing nonsensical, if sometimes hilarious, results" (p. 767). Clearly, Altavista Babel Fish did not match the sophistication of its literary namesake. The main concern from a teacher's perspective, however, seemed to be about its inappropriateness and inadequacy for language education, providing learners with "an easy way to dodge out of doing any work for an assignment and, in so doing, to avoid learning how to write in the target language" (p. 767).

A few decades later, more powerful and more accurate online translators, such as Google Translate, DeepL or Bing have replaced the now-extinct Altavista Babel Fish. Google Translate, the most well-known machine translation tool in 2022, provides real-time translation in over 100 languages. Translations for words, phrases or entire conversations can be entered by text, typed or finger-traced, or by voice, which in turn can be instantly transcribed. According to Google, the translator is used daily by over 500 million people, and as studies have shown, by numerous language students (Briggs, 2018; Delorme Benites *et al.*, 2021; Kok Wei, 2021), while language teachers still remain sceptical about their place in language education (Grove & Mundt, 2021).

Like many other machine translation tools, Google Translate started in 2006 as a statistical machine translation (SMT) service. Early studies reflect the shortcomings of SMT for language learning, often producing errors in word choice, syntax, and discourse (Jolley & Maimone, 2015; Niño, 2009; Williams, 2006). At the end of 2016, Google transitioned its translating method to neural machine translation (NMT). NMT uses deep learning techniques to translate whole sentences at a time, producing more accurate results. This development was a turning point in Machine Translation (MT). While previous MT systems stumbled over typos, reproducing the identical misspelt word in the translation, NMT does not only correct words. Drawing on a continuously expanding corpus (growing with every entry), it can translate less commonly used words, idioms, and colloquial language. Ducar and Schocket (2018) point out that Google Translate still fails to consider "context, connotation, denotation, register, and culture play in language production" (p. 785), and some issues with grammatical accuracy persist. Yet, as recent studies have shown, Google Translate outperforms the average intermediate-level language learner (Lee, 2022), leading some to claim that machine translation will be able to compete with human translation in the near future (Crossley, 2018, p. 543).

While Adam's Babel Fish will most certainly remain in the realm of utopia, the descendants of the AltaVista Babel Fish are likely to continue to challenge language education. Scholars are starting to realise that rather than fighting the use of machine translation tools teachers are better served by exploring how students can use them appropriately and adequately for language learning.

This chapter provides insights into the ways university students at different proficiency levels of language study in Chinese, French, German, Japanese and Spanish make use of machine translation tools. Contrary to the expectations of their teachers, many used the tools to support rather than circumvent learning, displaying innovative and creative learning strategies. The results of the study indicate that machine translation does not present a threat to language education, but that language teacher and

language learners need to develop MT literacy skills to evaluate and engage in new language learning opportunities in the 21<sup>st</sup> century.

### Literature Review

Language learners' frequent and extensive use of machine translation (MT) has been shown in several survey studies indicating MT is used in many aspects of language learning, both in and outside of class (Briggs, 2018; Clifford *et al.*, 2013; Jolley & Maimone, 2015;). Recent studies find that on average more than 90% of language students use MT, and more than half of them use them regularly (more than once a week) (Delorme Benites *et al.*, 2021; Kok Wei, 2021). Studies also document the students' recognition of practical benefits of MT, including fast and free access (Garcia & Pena, 2011), quick correction of spelling and grammar errors (Ducar & Schocket, 2018) and capability of multimodal inputs (Niño, 2020). In the context of second or foreign language (L2) learning, MT is commonly used for looking up the meaning of words for reading, translating L2 texts for comprehension, vocabulary learning, finding a translation of words in one's first language (L1) and checking grammar for writing, checking pronunciation for oral presentations (Briggs, 2018, Clifford *et al.*, 2013; Nugraha *et al.*, 2019; Kok Wei, 2021). Among these, using MT tools to look up words seems most popular among learners (Briggs, 2018; Clifford *et al.*, 2013). However, most of these studies are quantitative in nature. The details of how students use MT and how they perceive its role in language education have not been thoroughly investigated. Furthermore, recent studies involve the guided use of MT between specific language pairs, mostly among high proficiency learners. Thus, data is needed about how students of different languages at different levels of proficiency use MT.

Studies investigating language instructors' understanding of how their students use MT for L2 study indicate mixed perceptions, suggesting teachers are not entirely up to date regarding the availability and quality of MT tools and remain sceptical about the effectiveness of MT for language learning as well as their implications for academic integrity (Clifford *et al.*, 2013; Grove & Mundt, 2021; Jolley & Maimone, 2015; Lee, 2020; O'Neill, 2019; Vinall & Hellmich, 2021). For example, language instructors in Grove and Mundt's (2021) study expressed acceptance of using MTs to look up words for L2 reading but were reserved about the MT use for writing, especially above word level.

The benefits (and drawbacks) of using MT for L2 writing have been extensively investigated. Improvements in students' writing have been observed in terms of writing speed and fluency (Garcia & Pena, 2011; Ali & Alireza, 2014), accuracy (Chung & Ahn, 2021; Tsai, 2019), number of words written (Tsai, 2019), use of idiomatic language (Kol *et al.*, 2018), and overall writing quality (Lee, 2020). Although using MT as a dictionary in pre-and during-writing processes seems to be most popular among students (Clifford *et al.*, 2013; Lee, 2021), MT can also be used to check grammar such as conjugations, tenses, and other clues for contextual use of words (Briggs, 2018; Nugraha *et al.*, 2019). Among these different uses of MTs for writing, Ahn and Chung (2020) observed that high-proficiency students more frequently used MT for vocabulary search, while grammar checking was more common among low-proficiency students. Studies also suggest that MT is effective in writing revision (Clifford *et al.*, 2013; Briggs, 2018). These can be done by comparing students' drafts in L2 and MT outputs (Lee, 2020; Tsai, 2020) and by conducting reverse translation (Shei, 2002; Xu, 2020).

One of the most prominent issues in this process is the accuracy of MT outputs and how students manage them. Studies investigating the accuracy of MT outputs found differences in terms of translation units (e.g., word or phrase vs sentence or paragraph), MT types, text types and language pair (Lee, 2021). Although students seem to be generally satisfied with the accuracy of MT outputs (Delorme Benites *et al.*, 2021; Kok Wei, 2021), studies indicate that how students use MT outputs is significantly correlated with learners' proficiency levels. Higher-proficiency learners seem to be more successful

in detecting and correcting errors in MT outputs and using MT outputs to revise their L2 writing (Chung, 2020; Lee & Briggs, 2020). Observation of students' post-editing MT outputs revealed that lower-level students tend to focus on micro-level errors and lose sight of the larger context (Lee, 2017), while higher-proficiency students could effectively select better alternatives provided by MTs (Lee & Briggs, 2021). Chung and Ahn (2021) found low-proficiency students were more likely to accept MT outputs. Recent studies involving post-editing MT outputs report mostly positive writing quality outcomes and students' perceptions of MT (Chung & Ahn, 2021; Lee & Briggs, 2021; Tsai, 2019). Compared to research on students' proficiency levels, few studies investigate the differences associated with language pairs, and their findings seem sparse. Some studies examined the quality of Google Translate outputs and found differences in terms of intelligibility and accuracy rates favour alphabet-based languages compared to Asian languages (Aiken & Balan, 2011; Shadiev *et al.*, 2019). Other studies detected the reduced accuracy in Google Translate outputs between syntactically distant language pairs such as Arabic and English, Japanese and English and Persian and English (Afshin & Alaeddini, 2016; Correa, 2014; Hui, 2020; Shadiev *et al.*, 2019). Niño (2020), in her survey on MT effectiveness in conjunction with various task-based observations, also pointed out students' perception of MT accuracy depends on the language pair and the direction of translation (p. 9). However, her study is limited in that only two students studying non-alphabet-based languages (Japanese and Arabic) participated. At this point, how students of different languages handle the MT accuracy issues is not researched extensively (Lee, 2020).

Studies on students' perceptions about MT for L2 writing broadly indicate positive attitudes associating MT as a helpful and effective tool for L2 writing (Briggs, 2018; Chung & Ahn, 2021; Clifford *et al.*, 2013; Niño, 2009, 2020; O'Neill, 2019). Lee's (2021) meta-analysis of MT studies showed that students felt MT enabled them to better focus on the content, organisation, purpose, and conventions of their writing by aiding them with grammar and vocabulary. According to these studies, students reported that MT helped them increase their vocabulary, reduce grammatical errors, and gain confidence (Briggs, 2018; Lee, 2021). In addition, pedagogical benefits of MT include students' critical thinking about language use (Garcia, 2010), metalinguistic awareness of L2 (Correa, 2014), autonomous and self-directed approach to learning (Tsai, 2020) and increased confidence (Lee, 2020; Niño, 2009). However, some uncertainties about MT use among students have also been acknowledged. These include concerns for accuracy of outputs (Lee, 2019), a sense of guilt in using MTs (Xu, 2021) and anxiety about the lack of clear guidelines (Vinall & Hellmich, 2021). The perceptions about the usefulness and effectiveness of MTs among students seem to correlate with some learner variables. Garcia and Pena (2011) and Briggs (2018) indicated that low proficiency learners reported increased confidence in their writing associated with their MTs use, while other studies found that low proficiency learners considered MT as confusing and daunting (Niño, 2020; O'Neill, 2013; Stapleton & Kin, 2019).

Among teachers, concerns about the quality of MT, characteristic of Google Translate before it transitioned to the neural system in 2016, still seem persistent (Henshaw, 2020; Lee, 2020). Students presenting outputs from MT as their own writing is considered unethical among language instructors (Clifford *et al.*, 2013; Ducar & Schocket, 2018; Groves & Mundt, 2021). The assumption that students will develop a dependency on MT and use them indiscriminately is also a common concern, leading to the perception of MT as detrimental for language learning (Briggs, 2018; O'Neill, 2019).

At present, studies on students' practice and perceptions of MTs mainly focus either on a specific MT tool such as Google Translate (Briggs, 2018; Kok Wei, 2021) or investigate general MT use (Clifford *et al.*, 2013; Delorme Benites *et al.*, 2021). Lee (2020), in her meta-analysis of existing MT studies, points out that the majority of studies examine a single language pair, with English-Spanish being the most common. Since there are increasing numbers of language-specific MT tools available and their quality is continuously improving, investigating MT use for different language pairs and different



proficiency levels is needed to understand what students do and think about MT. This should contribute to more detailed and meaningful information that can fill the gap in the existing research, which will lead to well-informed policy development about MT use in L2 classrooms in the 21<sup>st</sup> century. This study examines the following four research questions:

1. In which L2 contexts do language learners use MT? Do MT practices differ between students of different languages and levels of proficiency? How do their experiences relate to teacher expectations of MT use?
2. How do language learners use MT to support their L2 writing? Do MT writing practices differ between students of different languages and levels of proficiency?
3. How do language learners manage MT output? Do post-editing practices differ among students of different languages and levels of proficiency?
4. To what extent is MT perceived as a helpful L2 writing tool by language teachers and students of different languages and different levels of proficiency?

## Method

This study uses a survey method to collect quantitative and qualitative data from language students and teachers at a tertiary institution in New Zealand. The preliminary analysis was presented in a previous publication (Alm & Watanabe, 2021).

### Instruments

Initial survey questions for students and teachers drew on Niño (2009), Clifford *et al.* (2013), Briggs (2018) and Tsai (2020). Four language students who had recently completed advanced-level language courses at the authors' institution were consulted to provide insights into their use of MT. The first version of the survey was pilot tested by five students who provided feedback. Since students were most familiar with the term Online Translator (OT), survey questions used OT to indicate machine translation. The final version of the student survey had 23 questions consisting of 12 single and multi-answer multiple-choice questions, seven Likert-scale questions and four open-ended questions. In addition, the multiple-choice questions had options for text entry for students to elaborate on their responses. Three sections of the survey were: 1) MT use for L2 learning in general, including frequency, types of devices, features of MT tools, and contexts of MT use (to assist reading, writing, speaking or listening); 2) MT use for L2 writing; 3) their perception about the helpfulness of MT.

The instructor survey contained 13 questions: three multi-answer questions, five open-ended questions and five Likert-scale questions. It was designed to collect data about their personal experience with MT, their perception of how students use MT for L2 study in general and L2 writing, their perception of the usefulness of MT. Some of the questions in the student survey were duplicated in the teacher survey with minor changes in wording to collect comparative data.

### Participants and Data Collection

An email invitation to participate in the online survey was sent to 438 students taking foreign language courses at a tertiary education institution in Semester 1, 2021. Their responses were collected for three weeks. In the email, students were informed that their participation was voluntary and that those who participated would go into the draw to win a supermarket voucher. We received 150 responses from students of all languages (Chinese, French, German, Japanese, and Spanish) and all proficiency levels (beginner, intermediate, and advanced). Nine students were taking two language courses and thus, the total number of

participants across five languages and three levels was 159 (Table 1). These students were not included in our quantitative data analysis, but their text entries were included in the thematic coding of the qualitative data. Languages were coded as CHIN (Chinese), FREN (French), GERM (German), JAPA (Japanese) and SPAN (Spanish). In terms of proficiency, we used levels of language courses and coded first-year courses as beginner (B), second year as intermediate (I) and third year as advanced (A). Since the instructional language of the tertiary institution in this study is English, all the participants were proficient English speakers.

Twelve out of 14 language teachers agreed to participate in the survey. Three Chinese, two French, two German, two Japanese, and three Spanish instructors participated. Five teach beginner-level courses, six and eight at intermediate and advanced levels, respectively. The languages and the teaching level of participating instructors are summarised in Table 2.

### Data Analysis

For this study, we selected survey questions relevant to our four research questions and applied quantitative and qualitative analysis. Quantitative data from survey questions were analysed by SPSS. Descriptive statistics of quantitative data were first generated for each question item, followed by a series of statistical tests. We used the following subcategories in terms of languages and proficiency levels to analyse student data.

**Table 1** Survey Participants' Language and Level: Students

Language	Respondents	Response rate
Chinese	14	24.56%
Japanese	36	36.00%
French	41	35.96%
German	29	59.18%
Spanish	39	33.05%
<b>Total</b>	<b>159</b>	<b>36.30%</b>
Beginner	56	27.32%
Intermediate	52	35.37%
Advanced	51	59.30%
<b>Total</b>	<b>159</b>	<b>36.30%</b>

**Table 2** Survey Participants' Language and Level: Teachers

Language	Levels	Respondents	Response rate
Chinese	B, I, A	3	100%
Japanese	B, I, A	2	100%
French	B, A	2	66.67%
German	B, I, A	2	100%
Spanish	B, I, A	3	75%
<b>Total</b>		<b>12</b>	<b>88.33%</b>

### **Language groups**

As discussed in the previous section, language pairs affect the quality of MT outputs (Goulet *et al.*, 2017; Hui, 2020; Niño, 2020), which could impact the way students use MT. As most of our participating students speak English as their L1, writing systems not based on the alphabet are likely to pose some challenges to using MT. Thus, to examine the differences in MT use between students of alphabet-based languages and non-alphabet-based languages, we set up Asian and alphabet-based subgroups. Chinese and Japanese are coded as “non-alphabet-based” (NAB) and French, German and Spanish as “alphabet-based” (AB). Participants of our survey consisted of 44 students studying “non-alphabet-based” languages (30.6%) and 100 studying “alphabet-based” languages (69.4%).

### **Language levels**

Of all the participants, 56 (35.22%) were enrolled at first-year level (beginner), 52 (32.70%) at second-year level (intermediate) and 51 (32.08%) at third-year level (advanced). Advanced-level students had the highest response rate (59.30%), followed by intermediate-level students (35.37%). For the statistical analysis, responses with two language choices (the nine students taking more than one language class at different levels) were deleted. For this reason, the number of language groups is 144 (44 non-alphabet-based, 100 alphabet-based), and 149 for language level (beginner 54, intermediate 50 and advanced 45).

### **Quantitative analysis**

For the analysis of multiple-choice questions, we used Pearson’s Chi-Square test. Chi-square test is appropriate for analysing categorical data, and both independent variables (e.g., teachers vs students and non-alphabet-based languages vs alphabet-based languages) and dependent variables (multiple-choice response on MT use) are categorical. Two types of statistical analysis were used for Likert-scale items. In the analysis of two categorical independent variable comparisons (teacher vs student and non-alphabet-based language vs alphabet-based language students) Mann-Whitney U test was applied to examine whether the distribution of the dependent variables is the same for the two groups. A Kruskal-Wallis test was used to determine whether the differences between the scores of beginners, intermediate and advanced learners were statistically significant. Further, a post hoc test was performed to test pairwise comparisons between language levels.

Results with  $p < 0.05$  were deemed statistically significant, meaning less than 5% probability that the null hypothesis is correct (i.e., no association). Thus, we reject the null and accept the alternative hypothesis. In this case, the alternative hypothesis is that there is an association between a type of person and response to these questions. To ensure the internal consistency of these multiple-item measurements, Cronbach’s alpha test was applied and considered scores of  $\alpha > 0.7$  as acceptable consistency except for one case, which showed  $\alpha = .6$  due to reverse coding.

### **Qualitative analysis**

To add depth to our quantitative results, the qualitative data from open-ended survey questions were analysed using content analysis (Dörnyei & Taguchi, 2009). The responses from clarification questions (comments of Likert-scale items) and short answer questions were analysed by coding responses according to the themes that emerged from the data in relation to our research questions. The first author created an initial list of codes, which was then revised and verified by the second author. We agreed on codes, categories and themes (collaboratively using Excel on OneDrive) and on quotations to illustrate our findings. Further, categories were quantified using frequency counts to create graphs.

## Results

### MT Practices

Language teachers guessed that 83% of their students would use MT for their language assignments. The student survey confirmed that 89.3% were using MT for their language study, with a higher percentage for advanced (95.6%) and intermediate students (94%) as opposed to 79.6% of students who started their language studies in that semester, presenting a significant difference as a chi-square test showed,  $\chi^2(2, N = 149) = 8.25, p = .016$ . The frequency of their MT use also increased with their proficiency level, but not significantly, with 23.8% of total participants using the tool daily, 56.4% a few times a week, and 19.8% less than once a week.

Depending on language and language level, students reported using different MT tools. Beginners of all languages almost exclusively used Google Translate. In contrast, students at higher proficiency levels complemented Google Translate with other translators, such as DeepL (German and French), or online dictionaries, such as Pleco (Chinese), Linguee (French, German), Word references and Reverso (French), Leo and Dict.cc (German), and SpanishDict (Spanish). Students at higher levels tended to use a wider variety of translation tools, in addition, or instead of Google Translate. A chi-square test revealed a significant association of the use of online dictionaries across proficiencies,  $\chi^2(2, N = 149) = 6.60, p = .037$ .

Students reported accessing MT on both phones and computers. Students of non-alphabet-based languages reported higher usage of phones (81.8%) compared to students of alphabet-based languages (61%), representing a significant difference,  $\chi^2(1) = 6.02, p = .014$ . Furthermore, the ability to use speech entries was mentioned as an advantage of phone use by Chinese and Japanese students, as “it is quicker than using the Japanese keyboard on my phone” (JAPA/B), and “helpful for characters that you can say but you don’t know the characters” (CHIN/A).

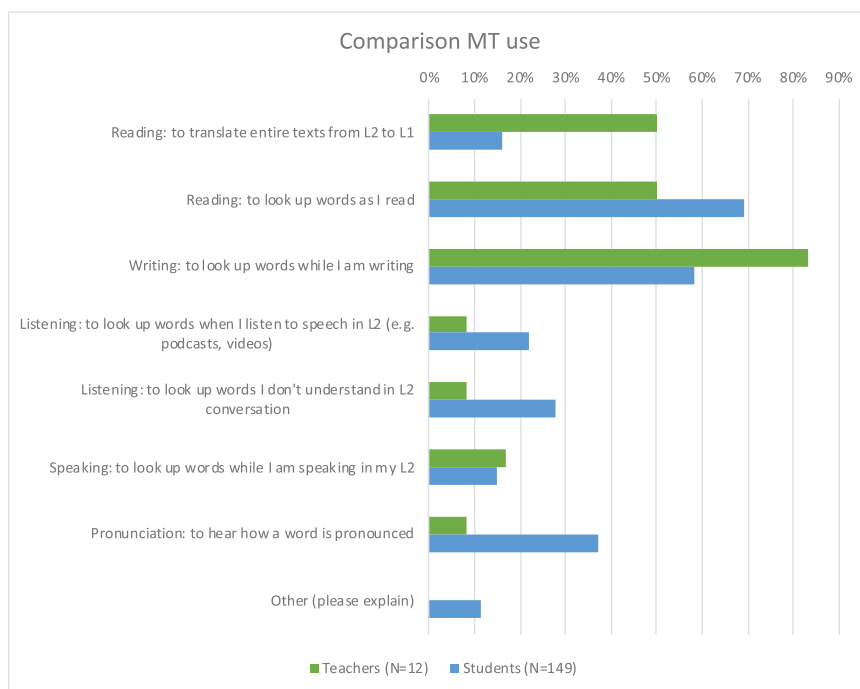
The preference for MT on computers increased with language proficiency (beginner 37%, intermediate 40%, and advanced 62.2%). The chi-square test confirmed a significant association between proficiency levels, with advanced learners being more strongly associated with MT on computers  $\chi^2(2, N = 149) = 7.23, p = .027$ . Advanced learners explained their preference with the ability to switch more easily between applications and the clearer layout of web-based applications.

### Comparison Student MT use and Teacher assumptions

Figure 1 displays the combined items of the student and teacher survey section on the reported use of MT in L2 contexts. The seven questions in multiple answer format were matched to compare teachers’ perceptions of students’ MT use and self-reported practices. Cronbach’s alphas for the teacher and student section were .75 and .66, respectively, indicating an acceptable level of reliability.

As depicted in Figure 1, students reported the use of MT for reading, not to translate entire texts, as anticipated by 50% of the teachers, but to look up words during reading. Teachers mainly expected students’ MT use for writing (83%), which was only the case for 58%. Students indicated MT use in listening contexts and to hear how words are pronounced (37%), a practice that only 8% of teachers envisaged. The chi-square tests showed significant differences between the teacher and student groups for the items on reading,  $\chi^2(1, N = 149) = 8.51, p = .004$ , and pronunciation,  $\chi^2(1, N = 149) = 3.94, p = .04$ .





**Figure 1** MT use in different contexts, Student and Teacher comparison.

### Comparison language groups

Twenty-six per cent of alphabet-based language students indicated using MT to look up words when listening to speech in L2, such as podcasts and videos, instead of only 9% students of non-alphabet-based languages, representing a statistically significant difference (Table 3). The qualitative data further suggests that the multimodal features of MT tools are mainly explored by students of non-alphabet-based languages, who reported using the photo and drawing features to input characters for translation, as explained by an advanced Japanese student: “If I am able to copy the kanji, I’ll paste it into Google Translate. If I can’t, I’ll write it on the drawing part of Google Translate, and sometimes I use the photo function and take a photo of the word or kanji” (JAPA/A).

### Comparison language levels

Advanced learners are more likely to use MT to look up words while reading (beginners 55.6%, intermediate 72%, advanced 80%) with learners at higher levels looking up words when writing. Advanced learners were most likely to use MT for other purposes, providing examples for looking up words, “to check words I can’t find in my dictionary” (SPAN/A), correcting grammar “to check the syntax” (JAPA/A) and verifying understanding “After reading a text to check how well I did” (JAPA/I).

**Table 3** Results for MT Use, Associations Between Groups

Item	Sub-groups	Value	df	p
Listening: to look up words when I listen to speech (e.g. podcasts, videos)	Language groups	5.28	1	.02*
Reading: to look up words as I read	Language levels	7.23	2	.027*
Writing: to look up words while I am writing	Language levels	4.53	2	.03*
Other	Language levels	6.84	2	.03*

\* $p < 0.05$

**Table 4** Results for MT Use in L2 Writing (Non-Alphabet-Based (NAB) and Alphabet-Based (AB languages))

Item	Group	N	Mean Rank	Sum of Ranks	U	P
1. To look up words (as a dictionary)	NAB	35	65.43	2290.00	1420.00	.457
	AB	88	60.64	5336.00		
2. To look up how an L2 word can be used in a sentence	NAB	35	74.00	2590.00	1120.000	<b>.014*</b>
	AB	88	57.23	5036.00		
3. To correct grammar	NAB	35	76.61	2681.50	1028.500	<b>.003*</b>
	AB	88	56.19	4944.50		
4. To make stylistic improvements to writing	NAB	35	71.61	2506.50	1203.500	.053
	AB	88	58.18	5119.50		

\* $p < 0.05$

### MT for L2 Writing

Most language students used MT to support their writing. However, MT writing practices differed among students of different languages and proficiency levels. The statistical analysis of the questionnaire section on MT use for L2 writing, consisting of four Likert-scale items from 1 (*extremely likely*) to 5 (*extremely unlikely*), Cronbach's alpha ( $N = 4$ ) .72, revealed significant differences of items 2 and 3 between students of non-alphabet-based and alphabet-based languages (Table 4).

The results indicate that students of all languages use MT to look up words. However, students of alphabet-based languages are more likely to use MT to improve their writing pragmatically, grammatically, and stylistically than students of non-alphabet-based languages.

To establish differences between language levels, a Kruskal-Willis test was applied. It revealed significant differences among the three proficiency level groups ( $H = 6.368$ ,  $p = .042$ ) for item 2. The post hoc tests to examine pairwise comparisons further showed that the advanced group was significantly different to the intermediate group ( $p = .047$ ) and the beginner group ( $p = .019$ ). The intermediate and beginner groups were not significantly different ( $p = .695$ ). This indicates that learners with higher proficiency levels are more likely to use MT to contextualise their word choices.

The qualitative data analysis provided further insights into L2 writing practices at different proficiency levels and between languages. Three practices could be identified: MT for 1) lexical searches, 2) grammar correction, and 3) proofreading.

### MT for lexical searches

Findings indicate that beginners are more likely to use MT to look up individual words, as a novice learner of German explained, "to increase my understanding and vocabulary" (GERM/B). Some beginners from both alphabet-based and non-alphabet-based languages justified their practice, stating that MTs "are not good enough to do long or complex sentences" (CHIN/B) or "they are not reliable for translating longer sentences" (FREN/B). With increasing language knowledge, learners explored MT for words in context (considering contextual features of a sentence) and phrases (idiomatic expressions). An advanced

learner of Japanese explained, “to see what synonyms the word has, and example sentences so I can see if it’s the right word to use (JAPA/A). The following example of an advanced student illustrates the use (and perceived capability) of MT for idiomatic language,

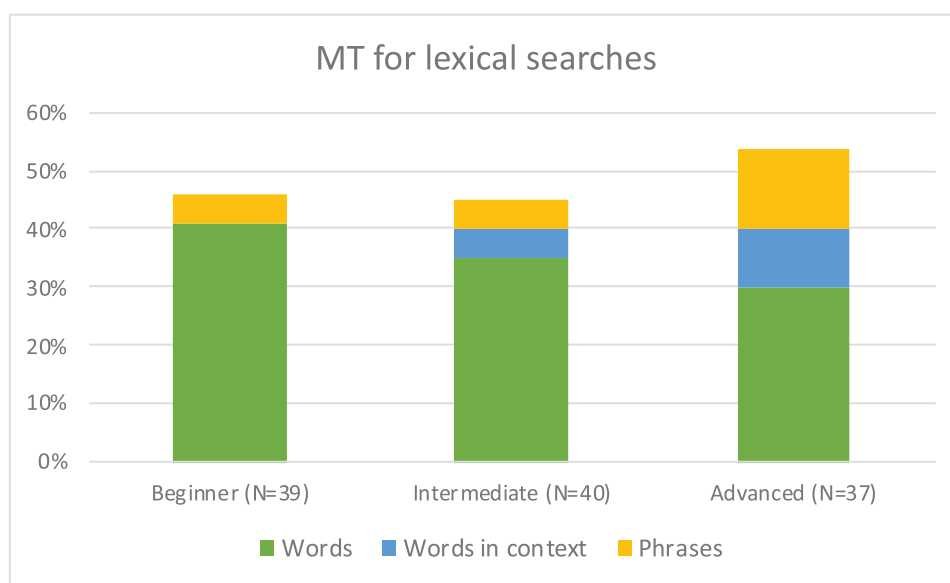
*Sometimes if I have a phrase in L1 that I am not sure if it is a widely used phrase across languages or an idiomatic phrase. IE ‘getting lost in my own thoughts’ this was a recent one I searched up to verify that it is also used in French. (FREN/A)*

Confirming the significant association of online dictionary use by advanced learners (noted earlier), advanced students of Spanish, French and German explained their use of SpanishDict, Reverso and Linguee to check “how a word should be used in a sentence” (SPAN/A), and “how people fit the phrase in a sentence” (FREN/A), to create their own sentences.

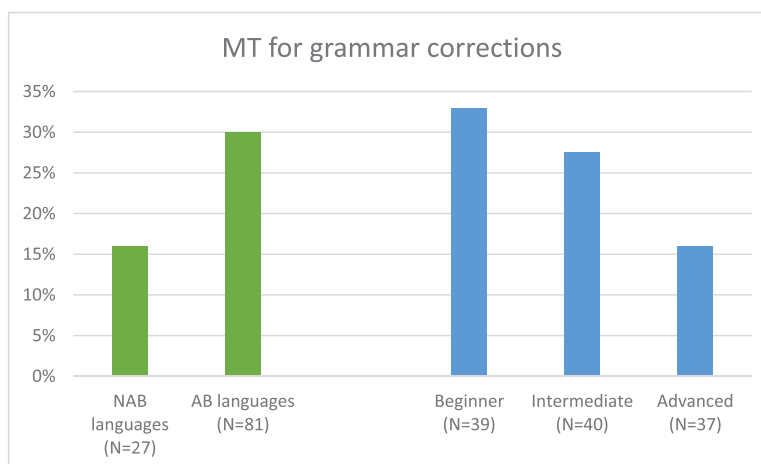
The analysis of MT for words, words in context, and phrases (Figure 2) shows a lower percentage of MT use at word level and a higher percentage of checking words in contexts and phrases by advanced learners. Proportionally more students of alphabet-based languages than non-alphabet-based languages (41% vs 32%) commented on the use of MT for lexical searches.

### **MT for grammar correction**

Students of alphabet-based languages provided more and a wider variety of examples for MT for grammar correction than students of non-alphabet-based languages (30% vs 16%). Beginners of all languages provided more examples for grammar applications (33%), followed by intermediate level learners (27.5%) and advanced learners (16.2%), as illustrated in Figure 3. Some alphabet-based language students used MT to check on conjugations (N = 12), as this intermediate Spanish student who used MT “To find the root word in L2 so that I can then conjugate it. Sometimes I use it to find out the conjugation especially if it’s an irregular verb” (SPAN/I), to correct syntactical errors (N = 7), “Since the structure of español sentences are very different to English, I sometimes get confused in which order the words go in, so I usually google the sentence to understand what order of the words” (SPAN/B) or to look up tenses, gender, adjective agreement, and accents.



**Figure 2** MT use for L2 writing compared by language level.



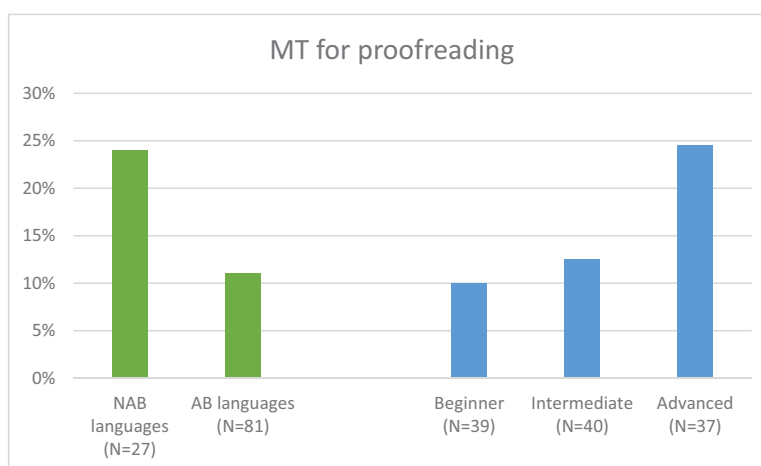
**Figure 3** *MT for L2 grammar correction compared by language and language level.*

### **MT for proofreading**

Some students used MT to proofread their texts. This practice was reported proportionally more often by students of non-alphabet-based languages (24% vs 11%), as Figure 4 shows. Many students used reverse translation (translating their L2 text back into their L1) to ensure their writing made sense. An advanced Spanish student explained, “Pasting whole sentences (I have already written in Spanish), to make sure they sound ok when translated back into English. (Usually, if the sentence sounds odd in English, then I might have made a mistake in the Spanish version)” (SPAN/A). A Chinese student describes the rationale for translating the translation back into the L2, “flicking back to L2 after converting to L1 makes a better sentence or introduces new words” (CHIN/B). Advanced students provided more examples of MT for proof-reading, such as this Japanese learner who would not use MT for grammatical improvement, “Google translate isn’t a great source of syntactical advice but if it does understand a complex sentence, it is likely that such a sentence is grammatical” (JAPA/A).

### **Dealing with MT Output**

Language learners differed in their assessment of MT output and their ability to manage it. The two items of the MT output section, *I trust the output of an OT for L2 writing* and *OTs are insufficiently accurate to be useful for L2 writing* (reversed) (Cronbach’s Alpha .58) on Likert-scale from 1 (*strongly agree*) to 5 (*strongly disagree*) show the differences between both language groups and proficiency levels.



**Figure 4** *MT for proofreading compared by language and language level.*

A Mann-Whitely test revealed that students of non-alphabet-based languages (Mdn = 3) were less trusting of MT output than students of alphabet-based languages (Mdn = 2), ( $U = 1132.500$ ,  $p = .022$ ). A Japanese student went as far as describing MT as “dodgy and inaccurate most of the time” (JAPA/I). Students of both non-alphabet-based and alphabet-based languages provided an interesting perspective. A third-year student of Japanese and German describes her experiences of MT with both languages:

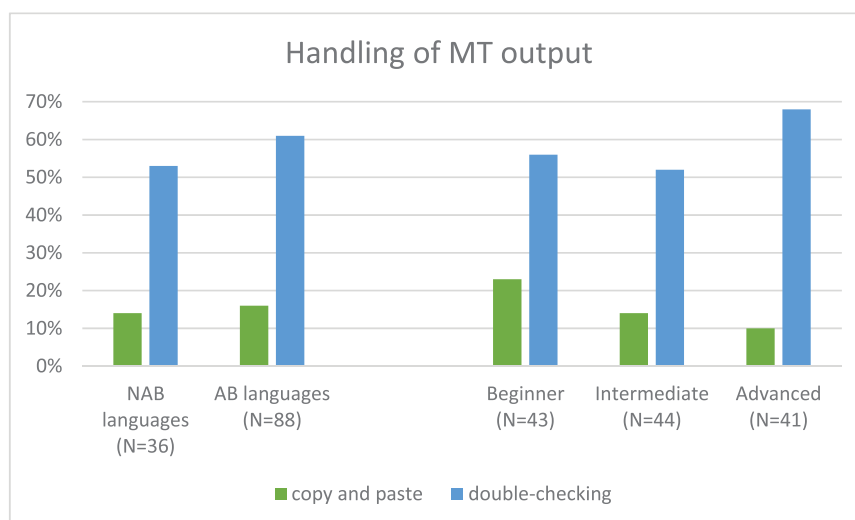
*For Japanese, when translating a long sentence, OT like google translate usually only gives an idea of what the sentence looks like and most likely the output sentence it is full of error or doesn't make any sense if the sentence is too complex and complicated ... However German-English on google translate are relatively accurate no matter the sentence is complex or not. (JAPA/A and GERM/A)*

She explained the difference between the different grammar systems of the languages, “Translate language with relatively well-structured grammar seems to be more accurate with OT. eg Ger-English. language like chinese/Jpn that have more flexible grammar items are harder to translate well with OTs” (JAPA/A and GERM/A).

The analysis of the second item, *OTs are insufficiently accurate to be useful for L2 writing*, showed marginal significance between languages ( $U = 1700.500$ ,  $p = .052$ ) but significant differences across proficiency levels. Beginners (Mdn = 3) are significantly more likely to say OTs are insufficiently accurate to be useful for L2 writing compared to advanced students (Mdn = 4), ( $H = 7.897$ ,  $p = .006$ ).

The qualitative data of the open-ended question on the students' handling of MT output (copy and paste vs double-checking) illustrate the impact of perceived MT quality on post-editing practices. Figure 5 illustrates the differences between languages and between proficiency levels.

Copy and paste practices decrease with developing proficiency, although it depends on specific writing contexts. Students reported using copy and paste mainly for single words. An advanced Spanish student admitted, “I copy and paste but I'll read the sentence and see if it looks right to me” (SPAN/A). A student of French and Japanese explained, “google translate (on browsers) has a function where common phrases' approved translations pop up automatically. I trust those” (FREN/A, JAPA/I). A Japanese student explained that they avoid copy and pasting because it “changes the format” (JAPA/I). Double-checking practices vary according to proficiency, with beginners seeking confirmation of accuracy primarily in



**Figure 5** Differences between languages and proficiency levels in managing MT output.



textbooks and class notes. In contrast, learners of higher proficiencies consult a wider range of language tools and apply strategies that involve their own knowledge, showing evidence that they interact with the output through reverse translation and input modification (Table 5 in Appendix A).

### MT as a Helpful L2 Writing Tool

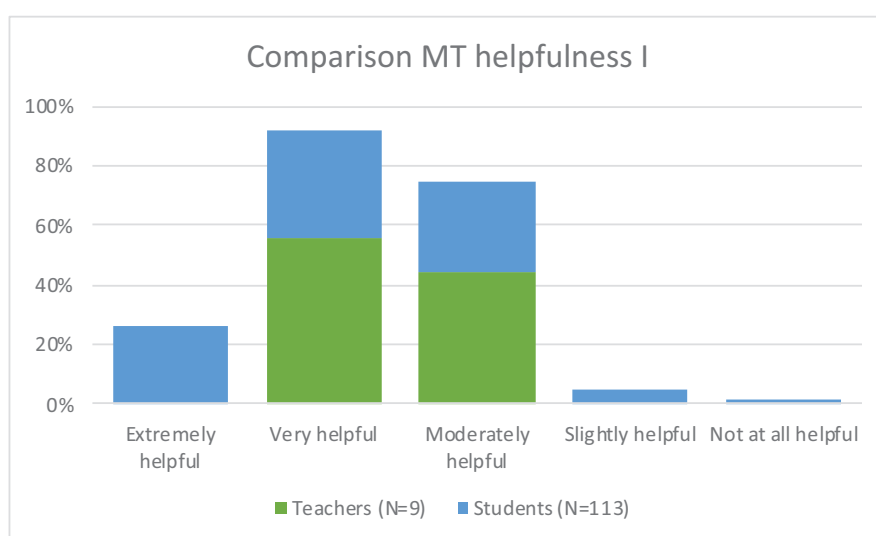
Language learners at all levels find MT helpful for L2 writing (Mdn = 2), and teachers recognise that their students find them helpful (Mdn = 2). Although the distribution over the Likert-scale from 1 (*extremely helpful*) to 5 (*not at all helpful*) differs, with a standard deviation of .53 for teachers, vs .95 for students (Figure 6), a Mann-Whitney test confirmed that there is no significant difference ( $U = 419.500, p = .358$ ) between student and teacher responses. The reasons for the perceived helpfulness, however, seems to differ. While teachers agree with the item *Students feel they do not need to learn to write in L2 because OTs can do the work* (Mdn = 2), students disagree strongly (Mdn = 5) and significantly ( $U = 1227.500, p = .001$ ) with the corresponding item.

Students also differed amongst subgroups in their responses to the question *OTs are helpful for L2 writing*. A Mann-Whitney test revealed a moderate association between languages. Alphabet-based language students are significantly more likely to say OTs are helpful for L2 writing compared with non-alphabet-based language students ( $U = 909.500, p = .043$ ), possibly because they trust more in their output.

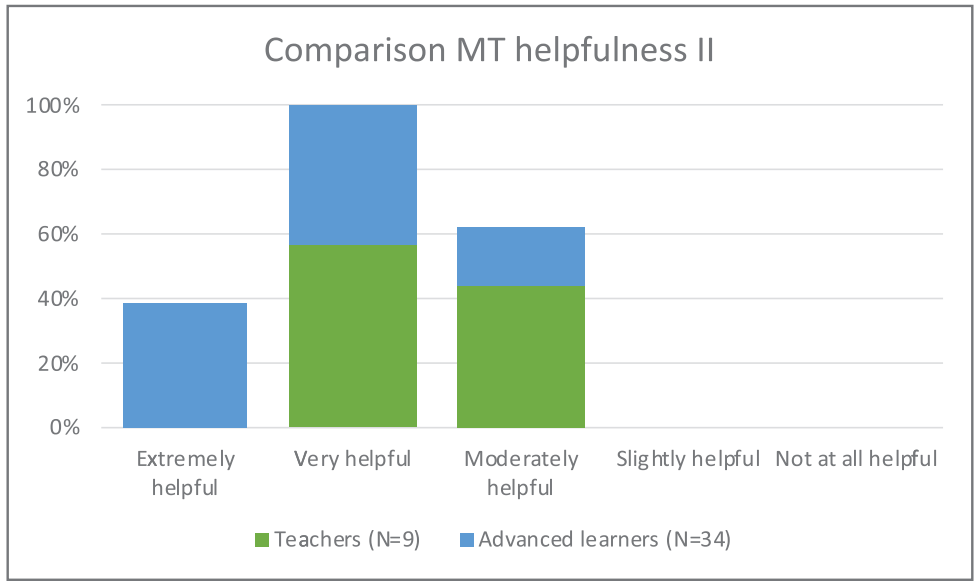
There is also a significant association between proficiency levels. A Kruskal-Wallis test indicated significant differences amongst language levels ( $H = 10.644, p = .005$ ). The pairwise comparisons further showed that the advanced group was significantly different to the intermediate group ( $p = .032$ ) and the beginner group ( $p = .001$ ).

The differences in perceived helpfulness led us to compare teachers and high proficiency learners, which showed a different picture from the initial student-teacher comparison (Figure 7). A Mann-Whitney test revealed that the differences between advanced learners and teachers are significant ( $U = 79.500, p = .018$ ).

Further, a Kruskal-Wallis test which compared the responses from teachers and all three proficiency groups confirmed significant differences among levels stated above ( $H = 12.001, p = .007$ ). The pairwise comparisons between advanced students and all groups are visually represented in Figure 8.



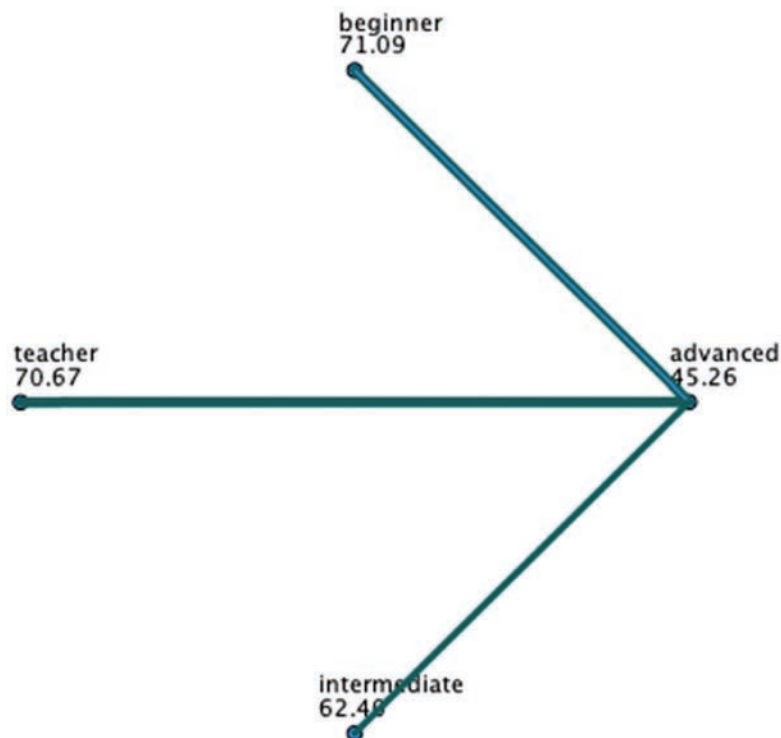
**Figure 6** Descriptive statistics: Teacher and student responses on perceived helpfulness of MT for L2 writing.



**Figure 7** Descriptive statistics: Teacher and advanced learner responses on perceived helpfulness of MT for L2 writing.

Each node shows the sample average rank of each group. Advanced-level students, who have the lowest mean rank (MR = 46.26), are significantly more likely to say OTs are helpful for L2 writing compared with beginners (MR = 71.09) and intermediate level students (MR = 62.40), as well as teachers (MR = 70.67).

The qualitative analysis established that high-proficiency learners find MT helpful because of its convenience and the guidance it provides in the writing process. They also address the need to use MT critically, which involves abstaining from bulk translation, being aware of MT dependency, and the need to assess



**Figure 8** Pairwise Comparisons of teachers and students: responses on perceived helpfulness of MT for L2 writing.

output (Table 6 in Appendix B). However, learners of lower levels of proficiency also provided valuable insights. Highlighting the convenience of MT, a beginner compared MT to calculators, “Much like using an online calculator to check maths questions” (SPAN/B). An intermediate learner of Chinese noted that MT tools provide guidance when a teacher is not available, or students seek to model their writing on native speaker examples, and therefore “contribute to the learning process.” Two intermediate French students wrote about the need to “interpret” and “critically assess” MT output. Some argued that a “certain level of language” is necessary to assess the output. Even beginners have provided examples of skilful MT use, such as breaking down sentences in recognisable chunks (GERM/B), indicating that they have their “brain switched on,” a crucial component of MT use, as pointed out by an advanced French student.

## Discussion

### The Different Contexts of MT Use (RQ1)

The results support studies that established frequent and extensive use of MT amongst language students (Briggs, 2018; Delorme Benites *et al.*, 2021; Kok Wei, 2021). We further found that MT use increases with language proficiency. More advanced learners also employ various MT tools, while beginners mainly rely on Google Translate. At higher proficiency levels, students tend to use MT tools in addition to Google Translate based on the corpora of their L2 and language-specific online dictionaries, which increasingly incorporate NMT.

Students use MT in a variety of L2 contexts. Contrary to teacher expectations that students (mis)use MT for bulk translation of L2 texts and in L2 writing, reflecting a traditional approach to classroom-based language learning based on reading and writing, students use MT in multimodal learning contexts and to make use of multimodal features. Similar to Niño (2020), we found that students use the sound output as an aid for pronunciation.

Students of French, German and Spanish reported using MT in multimodal learning contexts when listening to podcasts or viewing videos. This might be due to the higher availability of resources for learners of these languages or the higher accessibility of authentic resources for students of these alphabet-based languages, as they are deemed easier to learn and understand than Chinese or Japanese. Students of non-alphabet-based languages might also find it more difficult to spell an unknown word or phrase to look it up.

However, the multimodal features of MT tools are mainly explored by students of non-alphabet-based languages, who reported using the photo and drawing features on their phones to input characters for translation. Niño (2020) similarly detected a minority of students using MT for “picture translation of non-Latin script languages” in an informal setting (p. 7). Greater use of phones among non-alphabet-based language students reveals the challenges of the non-alphabetical writing system of Chinese and Japanese, which might be overcome using flexible input methods such as voice recognition, image capture, and the drawing pad.

### MT for L2 Writing (RQ2)

For most language learners, MT tools replace conventional dictionaries during the writing process. Our results confirm observations of pre-and-post-NMT studies (Clifford *et al.*, 2013; Lee, 2021) which all found that students use MT to look up words, a practice also endorsed by most teachers. Like Briggs (2018), we established that students use MT to improve their writing pragmatically and grammatically. We found that learners with higher proficiency levels are more likely to use MT to contextualise their word choices, a practice also shown in Ahn and Chung (2020) and Ryu *et al.* (2022). Furthermore,

MT use for grammatical corrections seems particularly appealing for learners at lower proficiency levels, as Ahn and Chung (2020) observed. Relating to alphabet-based languages (French and Spanish), Ducar and Schocket (2018) point out that “GT is so good at conjugating that it often allows lower-level students to produce complicated verb tenses that have not yet been studied” (pp. 782–783). In contrast, for non-alphabet-based languages, Lee and Briggs (2021) found that Korean English learners of lower proficiency levels used MT to correct minor errors associated with definite and indefinite articles and placement of prepositions while giving up on correcting more complex global errors. The reluctance to resort to MT for grammatical corrections expressed mainly by Japanese students in our study might be explained by the difference in syntax between English and non-alphabet-based languages, which could cause extra challenges for non-alphabet-based language learners. Hui (2020) found issues in distinguishing written and spoken styles (affects the choice of plain/polite forms), translating sentences with missing subject or object, detecting contextual switching of pronouns and proper nouns and idiomatic expressions. These language features permeate all levels of Japanese writing and frequently pose challenges to English learners of Japanese.

MT for proofreading emerged in our study as a widespread practice, reported proportionally more often by students of non-alphabet-based languages and mainly by students at higher levels. Our participants reported using MT to compare their L2 drafts with MT output. Their developed language skills allow high-proficiency learners, as noted by Lee and Briggs (2021), to “notice key differences between their L2 texts and MT outputs and to use that information to effectively detect and treat errors in their L2 composition” (p. 29). Many students also used reverse translation as a way to ensure their writing made sense or to improve it. Xu (2021) observed that students repeated these processes until they had acceptable results. In these cases, Lee (2021) suggests that MT could function “similar to peer-editing” (p. 168) at least at word or sentence level by generating immediate and personalised feedback.

### **MT Output and the Ability to Manage It (RQ3)**

The accuracy of MT output depends on several factors, including the MT tool itself, the suitability of the tool for a specific language pair, the source text, and the input units (words as opposed to phrases or bulk texts) (Lee, 2021). Our participants differed in their assessment of MT output and their ability to manage it. Students of non-alphabet-based languages were generally less trusting of MT and explained their position with the more complex grammatical system of non-alphabet-based languages. They seemed less reliant on MT for this reason. However, recent studies on syntactically distant language pairs, such as Korean and English, indicate that the accuracy of MT output is continuously improving (Lee, 2022).

Low-proficiency learners were generally also more critical of the accuracy of MT output. Their position, however, is likely to relate to their limited language skills and inability to discriminate correct from erroneous output. In addition, their understanding of MT seems to be informed by pre-NMT. They accept (and copy) translations at the word level yet avoid output at sentence level, which, in fact, produces more accurate results with NMT (Ducar & Schocket, 2018).

However, learners at higher levels of proficiency are less critical of MT output and seem to have a better understanding of how they work. They are also better equipped to detect and correct errors in the output and learn from the output that offers new lexical and grammatical information (cf. Chung, 2020; Lee & Briggs, 2020). In addition, our participants commented on a variety of strategies to double-check MT output with other tools and resources. Further, the use of repeated reverse translation and input manipulation indicates that advanced learners “interact with the machine,” as Bowker (2021, p. 28) put it. Examples provided by students, such as chunking and rewording (see examples in Table 5 in Appendix A), indicate that some learners are able to combine their advanced understanding of the L2 with their understanding of how MT works, demonstrating MT literacy.

## MT as a Helpful Tool for L2 Writing

MT literacy has been defined by Bowker (2020) as the ability to use MT critically. As she puts it, “Using machine translation is easy. Using it *critically* requires thought” (p. 28). Both aspects are represented in participants’ responses on MT helpfulness for L2 writing. The perception of MT as being easy is shared by most. Students appreciate the convenience of an accessible language tool. Teachers, however, seem to be of the view that MT makes L2 writing too easy for language learners, anticipating negative impacts on language learning development, MT dependency (cf. O’Neill, 2019) and fearing issues of academic integrity (cf. Groves & Mundt, 2021), and the undermining effect of their own authority (cf. Vinall & Hellmich, 2021).

Student responses, however, provide a different picture. The statistical arrow in Figure 8, headed by advanced language learners, takes symbolic meaning, suggesting that this group of students is leading the way for critical MT use. It does not seem to be a coincidence that advanced-level students present the highest response rate (59.30%) of the survey. Their comments on MT helpfulness for L2 writing provide valuable insights into the awareness of students of MT in supporting language learning development, as MT is not only easy to use but also offers constructive guidance during the writing process, which often takes place in the absence of a teacher, leading in many cases to lexical and grammatical improvement (cf. Briggs, 2018), enhancement of metalinguistic awareness (Correa, 2014), and as Tsai (2020) found, support of autonomous learning. Advanced students in particular seem cognisant of the detrimental effect of MT use for bulk translation and MT dependency on language development and the need to assess outputs critically.

The insights gained from student responses suggest that the concerns of many teachers about negative impacts of MT use on language learning, and fears about their status in the language classroom are largely unsubstantiated. Rather, they highlight the urgent need for MT literacy development for both teachers and students, possibly headed by advanced and experienced language learners, paving the way to new pedagogies in language learning.

## Conclusion

In 2010, Garcia entitled his article on the benefits of the Google Translator Toolkit for translation *Is machine translation ready yet?* Evoking the utopian dream of a universal translator, Garcia, in fact, referred to the readiness of Google Translate for machine translation post-editing (MTPE), a translation technique in which machine translation serves as the basis for subsequent editing by (human) translators. Similarly, earlier studies have shown that inaccurate MT can support language learning, as it requires learners to actively engage with the output, supporting the development of meta-cognitive language skills (Correa, 2014). With the emergence of highly accurate machine translation services, such as Google Translate and DeepL, the question of readiness seems to target the humans rather than the machine. Are translators, teachers, and students ready to adapt and develop the skills to use MT critically and allow MT to be powerful guides in the language learning process?

From our perspective, MT literacy includes an understanding of the functional aspects of MT. Only a few students from our cohort took advantage of the multimodal features Google Translate provides. The practicality of the voice input for learners of non-alphabet-based languages or the ability to draw characters instead of typing them are just two examples that illustrate how the technology can extend traditional learning practices. Technical features become affordances for language learning as learners discover how they can use these features to suit their personal learning contexts.

Our study has also shown that many teachers and students are still stuck in pre-NMT times and don’t necessarily understand how Google Translate, DeepL and Naver Papago differ from online dictionaries. An understanding of the systems that power MT, working on sentence rather than word level, is another technical aspect of MT literacy.



Several studies report on the benefits of MT training (Faber & Turrero-Garcia, 2020; Hellmich, 2020; O'Neill, 2019) and on the development of students' MT literacy. The findings of our study suggest that teachers also need to upskill and be trained in functional and critical MT use. As effective users of MT, they will be able to design language learning activities using the strength of MT, suitable for language learners of different languages and at different levels of proficiency.

## References

- Adams, D. *The Hitchhiker's Guide to the Galaxy*. New York: Harmony Books, 1979.
- Afshin, H., & Alaeddini, M. A. (2016). A contrastive analysis of machine translation (Google Translate) and human translation: Efficacy in translating verb tense from English to Persian. *Mediterranean Journal of Social Sciences*, 7(4), 40–48. <https://doi.org/10.5901/mjss.2016.v7n4S2p40>
- 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.
- Aiken, M., & Balan, S. (2011). An analysis of Google Translate accuracy. *Translation Journal*, 16(2). <http://translationjournal.net/journal/56google.htm>
- 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>
- Ali, K., & Alireza, F. (2014). The effect of computer-assisted translation on L2 learners' mastery of writing. *International Journal of Research Studies in Language Learning*, 3, 29–44.
- Alm, A., & Watanabe, Y. (2021). Online translators for L2 writing: A comparison of student and teacher perspectives. In N. Zoghلامي, C. Brudermann, C. Sarré, M. Grosbois, L. Bradley & S. Thouëсны (Eds.), *CALL and Professionalisation: Short Papers from EUROCALL 2021*. (pp. 23–28). Research-publishing.net. <https://doi.org/10.14705/rpnet.2021.54.1303>
- Bowker, L. (2020). Machine translation literacy instruction for international business students and business English instructors. *Journal of Business & Finance Librarianship*, 25(1–2), 25–43. <https://doi.org/10.1080/08963568.2020.1794739>
- Bowker, L. (2021). Machine translation use outside the language industries: A comparison of five delivery formats for machine translation literacy instruction. *Translation and Interpreting Technology Online*, 25–36. [https://doi.org/10.26615/978-954-452-071-7\\_004](https://doi.org/10.26615/978-954-452-071-7_004)
- Briggs, N. (2018). Neural machine translation tools in the language learning classroom: Students' use, perceptions, and analyses. *The JALT CALL Journal*, 14(1), 3–24. <https://doi.org/10.29140/jalt-call.v14n1.221>
- Chung, E. S. (2020). The effect of L2 proficiency on post-editing machine translated texts. *The Journal of Asia TEFL*, 17(1), 182–193. <https://doi.org/10.18823/asiatefl.2020.17.1.11.182>
- Chung, E. S., & Ahn, S. (2021). The effect of using machine translation on linguistic features in L2 writing across proficiency levels and text genres. *Computer Assisted Language Learning*, 1–26. Advance online publication. <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? *@tic. Revista d'innovació Educativa*, (10), 108–121. <https://doi.org/10.7203/attic.10.2228>
- 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 and Language Integrated Learning*, 7(1), 1–20. <https://doi.org/10.5294/laclil.2014.7.1.1>
- Crossely, S. (2018). Technological disruption in foreign language teaching: The rise of simultaneous machine translation. *Language Teaching*, 51(4), 541–552. <https://doi.org/10.1017/S0261444818000253>
- Delorme Benites, A., Cotelli Kureth, S., Lehr, C., & Steele, E. (2021). Machine translation literacy: A panorama of practices at Swiss universities and implications for language teaching. In N. Zoghلامي, C. Brudermann, C. Sarré, M. Grosbois, L. Bradley, & S. Thouëсны (Eds.), *CALL and*

- professionalisation: Short papers from EUROCALL 2021 (pp. 80–87). Research-publishing.net. <https://doi.org/10.14705/rpnet.2021.54.1313>
- Dörnyei, Z., & Taguchi, T. (2009). *Questionnaires in second language research: Construction, administration, and processing*. Routledge.
- Ducar, C., & Schocket, D. H. (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>
- Faber, A., & Turrero-Garcia, M. (2020, March 10). Online translators as a pedagogical tool. *The FLT MAG*. Retrieved from <https://fltmag.com>
- Farzi, R. (2016). *Taming translation technology for L2 writing: Documenting the use of free online translation tools by ESL students in a writing course* [Doctoral thesis, University of Ottawa]. uO Research. Retrieved from <http://www.ruor.uottawa.ca/handle/10393/34585>
- Garcia, I. (2010). Is machine translation ready yet? *Target. International Journal of Translation Studies*, 22(1), 7–21. <https://doi.org/10.1075/target.22.1.02gar>
- 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>
- Giannetti, T. R. (2016). *Google translate as a resource for writing: A study of error production in seventh grade Spanish*. [Master's thesis, St. John Fisher College]. Fisher Digital Publications. [https://fisherpub.sjfc.edu/education\\_ETD\\_masters/353](https://fisherpub.sjfc.edu/education_ETD_masters/353)
- Groves, M., & Mundt, K. (2021). A ghostwriter in the machine? Attitudes of academic staff towards machine translation use in internationalised higher education. *Journal of English for Academic Purposes*, 50, 1–11. <https://doi.org/10.1016/j.jeap.2021.100957>
- Hellmich, E. A. (2021). Machine translation in foreign language writing: Student use to guide pedagogical practice. *Alsic. Apprentissage Des Langues et Systèmes d'Information et de Communication*, 24(1), 1–15. <https://doi.org/10.4000/alsic.5705>
- Henshaw, F. (2020, June 15). Online translators in language classes: Pedagogical and practical considerations. *The FLT MAG*. <https://fltmag.com>
- Hui, H. (2020). A study on the language expressions and the pragmatic problems of machine translation: From a viewpoint of elementary Japanese language learners. *Bulletin of Tokyo Gakugei University*, 20, 557–565.
- 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). Richmond: Robert M. Terry.
- Kok Wei, L. (2021). The use of Google Translate in English language learning: How students view it. *International Journal of Advanced Research in Education and Society*, 3(1), 47–53. <https://myjms.mohe.gov.my/index.php/ijares/article/view/12459>
- Kol, S., Schcolnik, M., & Spector-Cohen, E. (2018). Google Translate in academic writing courses? *The EuroCALL Review*, 26(2), 50–57. <https://doi.org/10.4995/eurocall.2018.10140>
- 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. (2021). The effectiveness of machine translation in foreign language education: A systematic review and meta-analysis. *Computer Assisted Language Learning*, 1–23. Advance online publication. <https://doi.org/10.1080/09588221.2021.1901745>
- Lee, S.-M. (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.-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>
- Luton, L. (2003). If the computer did my homework, how come I didn't get an "A"? *The French Review*, 766–770.

- Niño, A. (2009). Machine translation in foreign language learning: Language learners' and tutors' perceptions of its advantages and disadvantages. *ReCALL*, 21(2), 241–258. <https://doi.org/10.1017/S0958344009000172>
- Niño, A. (2020). Exploring the use of online machine translation for independent language learning. *Research in Learning Technology*, 28. <https://doi.org/10.25304/rlt.v28.2402>
- Nugraha, G., Ratnawati, R., & Surachmat, A. (2019). Exploring low and high students' perception on engaging e-dictionary in mastering vocabulary: Cross-sectional survey. *Indonesian EFL Journal*, 5(1), 37–44. <https://doi.org/10.25134/ieflj.v5i1.1609>
- O'Neill, E. M. (2012). *The effect of online translators on L2 writing in French* [Doctoral dissertation, University of Illinois at Urbana-Champaign]. ProQuest Dissertations Publishing. <https://www.proquest.com/docview/1447020851/abstract/85E12B4A363145DEPQ/1>
- O'Neill, E. M. (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(2), 47–65. <https://doi.org/10.5861/ijrsl.2019.4002>
- Ryu, J., Kim, Y.A., Park, S., Eum, S., Chun, S., & Yang, S. (2022). Exploring foreign language students' perceptions of the Guided Use of Machine Translation (GUMT) model for Korean writing. *L2 Journal*, 14(1), 136–165.
- Shadiev, R., Sun, A., & Huang, Y.-M. (2019). 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>
- Shei, C.-C. (2002). Teaching MT through pre-editing: Three case studies. *Proceedings of the 6th EAMT Workshop: Teaching Machine Translation*. Retrieved from <https://aclanthology.org/2002.eamt-1.10.pdf>
- Stapleton, P., & Kin, B. L. K. (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>
- Tsai, S.-C. (2020). Chinese students' perceptions of using Google Translate as a translingual CALL tool in EFL writing. *Computer Assisted Language Learning*, 1–23. <https://doi.org/10.1080/09588221.2020.1799412>
- Vinall, K., & Hellmich, E. A. (2021). Down the rabbit hole: Machine translation, metaphor, and instructor identity and agency. *Second Language Research & Practice*, 2(1), 99–118. <https://doi.org/10.125/69860>
- Williams, L. (2006). Web-based Machine Translation as a tool for promoting electronic literacy and language awareness. *Foreign Language Annals*, 39(4), 565–578. <https://doi.org/10.1111/j.1944-9720.2006.tb02276.x>
- Wuttikrikunlaya, P., Singhasiri, W., & Keyuravong, S. (2018). The use of online tools in L2 writing: A study of Thai university students. *Journal of English Language Teaching and English Linguistics*, 30(1), 107–148.
- Xu, J. (2020). Machine translation for editing compositions in a Chinese language class: Task design and student beliefs. *Journal of Technology and Chinese Language Teaching*, 11(1), 1–18.
- Xu, J. (2021). Google Translate for writing in a Japanese class: What students do and think. *The Journal of the National Council of Less Commonly Taught Languages*, 30, 136–182.

### Author Biodata

**Antonie Alm** (PhD, UCLA) is an associate professor at the University of Otago, New Zealand where she teaches German, Computer Assisted Language Learning and Intercultural Communication. Her research interests lie in L2 motivation and engagement, learner autonomy, informal language learning, and more recently the use of machine translation in language education. Antonie is on several editorial

review committees and works as an associate editor for the CALL Journal and the JALT CALL Journal. She is also the senior vice-president of the New Zealand German teacher association.

**Yuki Watanabe** (PhD, University of Kansas) is a lecturer at the University of Otago, New Zealand. She teaches courses in Communication, Sociology, Gender Studies, and Japanese. Yuki's research focuses on ways in which electronically mediated communication of specific types can enable meaningful social interaction. Her current stream of research has a sociological focus and draws on multi-disciplinary methods to concentrate on three themes: (i) media studies techniques for analysing and understanding contemporary popular culture in the global context (ii) the intersection of technology and intercultural communication, and (iii) communication technology applied in the domain of education and learning outcomes.

## Appendixes

### Appendix A

**Table 5** *Qualitative Analysis: Double-checking MT Output*

Category	Code and count	Example
People	Native speakers (4)	friends who are Chinese (CHIN/A) a native speaker on an online forum like reddit or duolingo forums (FREN/I) Spanish speaking husband (SPAN/B)
	Peers (3)	asking friends in class (JAPA/I)
	Teacher (3)	email lecturer (JAPA/A)
Tools	Online dictionaries (11)	other sites e.g. LineDict (CHIN/A) using a context translator (e.g. Reverso context) (FREN/A)
	Other OTs (6)	using 2 different OTs for clarification (FREN/A)
	Word	typing in into the Word document (FREN/A)
	Other sites, incl. Google (9)	Google it (SPAN/A)
Resources	Class materials/ textbook/ notes (20)	looking at notes from class or in textbook (SPAN/B)
Strategies	Own knowledge (17)	modify the outputs to fit what I already know (GERM/I)
	Reverse translation (13)	swap the language a couple times once I've written something (GERM/A)
	Changing input (4)	break down sentences into smaller units (SPAN/I)
		changing the context around the desired word to see how it changes (FREN/I) reword the L1 sentence a few times to make sure the context of the translation is correct (GERM/A)
Listening to output	putting smaller phrases from a bigger phrase and seeing if they are translated the same (SPAN/I) listening to it (JAPA/A)	



## Appendix B

Table 6 *Qualitative Data for Perceived Usefulness of MT*

Category	Example
Faster than dictionary	I think without them I would find it a lot harder to complete writing assignments as they are so quick and easy to use, rather than having to use a dictionary which is extremely time consuming (GERM/A).
Provides guidance	Yes I do, often we don't have the time to go to the office hours and ask the tutors for help ... (CHIN/A). Helps improve the accuracy of writing for learning purposes – and helps students to learn more about writing (FREN/A). It helps us to learn to write like a native speaker, produce writing that flows better (FREN/A). ... used as a resource in conjunction with a base writing ability in the language, using OTs provides a lot of room for a greater development of writing confidence and vocabulary (JAPA/A, SPAN/A).
Avoid bulk translation	Helps with specific words or grammar points but should not be used for whole sentences or paragraphs otherwise you are hindering your learning (FREN/A) I would never use google translate or 'bulk translate' a passage. I trust dictionaries more than a bulk translator (FREN/A).
Avoid OT dependency	They're great, but it's important not to rely on them.... Useful in moderation! Better not to have access to them in test conditions, as it challenges you more to use your own brain (FREN/A) It's up to the individual student ... to monitor their own reliance on OT or lack thereof and the effect that that has on their language learning (FREN/A, JAPA/I).
Need to interpret / assess output	You can only really use these websites if you have a certain level of language to begin with. So, they are helpful for learning new vocabulary etc. but you still have to have your brain switched on (FREN/A) I find that OTs work successfully in conjunction with an individual's existing foundational writing ability and can offer the opportunity to expand one's vocabulary (JAPA/A, SPAN/A). For my french level now, I would rather switch to google.fr (French google) and search up a topic related to what it is I am writing about to see how native speakers have gone about writing on the topic (FREN/A).